# AJS

(15)

Cole - H05457 - 15 - Icp787

## CONTENTS

# IN THIS ISSUE

LINCOLN QUILLIAN is assistant professer of sociology at the University of Wisconsin, Madison. He is interested in social stratification, race and ethnic relations, and statistics. He is currently working on several projects on the nature of causes of change over time in racial and class segregation in American cities.

PAMELA E. OLIVER is professor of sociology at the University of Wisconsin, Madison and past chair of the Collective Behavior and Social Movements section of the ASA. She continues to research media coverage of protest events and diffusion processes in protest cycles.

DANIEL J. MYERS is assistant professor of sociology at the University of Notre Dame. His current research focuses on the diffusion of collective action, behavioral game theory, and race-related rioting.

PAMELA PAXTON is assistant professor of sociology at Ohio State University. Her research interests include political sociology, quantitative methods, and stratification. Currently, she is investigating the relationship between social capital and democracy cross-nationality.

MICHAEL SUK-YOUNG CHWE is assistant professor of economics at the University of Chicago. He is currently writing a book manuscript entitled *Rational Ritual: Culture, Coordination, and Common Knowledge,* which uses rational-choice reasoning to understand cultural practices such as public rituals, political rallies, and television advertising.

DAVID JACOBS, professor of sociology at Ohio State University, does research on the politics of crime and social control, issues in political economy, and social organization. Current projects include studies of outcomes in the criminal justice system and longitudinal studies of political and legal change.

KATHERINE WOOD is currently a law student at the University of Chicago Law School.

SHIN-KAP HAN is assistant professor of sociology at Cornell University. He is currently working on two projects: one on coupled careers and the other on the network structure of the decision-making process in the U.S. Supreme Court.

PHYLLIS MOEN is the Ferris Family Professor of Life Course Studies as well as professor of sociology and human development at Cornell University. Her research focuses on work/family and life course transitions. She directs the Cornell Employment and Family Careers Institute, funded by the Alfred P. Sloan Foundation and codirects the Cornell Gerontology Research Institute, funded by the National Institute on Aging.

## Information for Contributors

**Editorial Procedures:** All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. Manuscripts are accepted for publication subject to the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal.* No paper known to be under consideration by any other journal will be reviewed by the *AJS*. In addition, the *AJS* is not a suitable vehicle for materials published elsewhere, in whole or in part. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his or her article without charge in any book of which he or she is the author or editor after it has appeared in the *Journal.*

**Manuscript Acceptance Policy:** While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the *Journal* has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the *Journal* in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

### Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, tables, and references—*double spaced,* allowing generous margins at the top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Do not use proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, beginning with the acknowledgment note as number 1. *Provide a word count on the cover page of the manuscript.*

2. Begin each table on a separate page. Refer to each table in numerical order in the text. Tables should not contain vertical lines or more than 20 two-digit columns or the equivalent.

3. Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.

4. Original or glossy prints of professionally drawn figures may be required if manuscript is accepted.

5. Include a brief abstract (not more than 100 words) summarizing the findings.

6. Submit *five copies* of the manuscript and abstract on white paper. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose a self-addressed stamped postcard to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article submitted to the *AJS* must be accompanied by a check or money order for $15.00, payable to the University of Chicago Press in U.S. currency or

its equivalent. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status. Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

**Format of References in Text:** All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61–65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of the Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (*a, b*) attached to the year of publication: (1965*a*).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons: (Duncan 1959; Gouldner 1963; Lipset 1964, pp. 61–65).

6. Format for references to machine-readable data files is

Davis, James Allan. 1978. GSS (General Social Survey). NORC ed. Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

**Format of References:** List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "REFERENCES." For examples, see text of articles in this issue.

(Rev. 3/98)

## CONTENTS

## IN THIS ISSUE

TIMOTHY J. BIBLARZ is assistant professer of sociology and research associate in the Andrus Gerontology Center at the University of Southern California. His research focuses on relationships between the family and patterns of social inequality, with an eye toward intergenerational issues. Current projects include a study of occupational transmission from mothers and fathers to sons and daughters, and an investigation of historical change in the relationship between family structure and children's educational transitions, from middle school to graduate school.

ADRIAN E. RAFTERY is professor of statistics and sociology at the University of Washington. He is currently editor of *Journal of the American Statistical Association* and was editor of *Sociological Methodology* during 1996–98. He is currently working on Bayesian model selection, on model-based cluster analysis, and on inference for population dynamics models.

JOHN BELLAMY FOSTER is associate professor of sociology at the University of Oregon and coeditor of *Organization and Environment* (Sage). He wrote the introduction to a new edition of Harr Braverman's *Labor and Monopoly Capital* (Monthly Review Press, 1998).

LAUREN B. EDELMAN is professor of law and sociology at the University of California, Berkeley, in the Jurisprudence and Social Policy Program and the Center for the Study of Law and Society. Her work addresses the relationship of organizations to their legal environments, organizational dispute resolution, and the internal legal cultures of organizations. She is currently coediting the law section of the *International Encyclopedia of the Social and Behavioral Sciences*.

CHRISTOPHER UGGEN is assistant professor of sociology at the University of Minnesota. His research addresses crime, law, and the employment relation. He is currently examining the socioeconomic determinants of illegal earnings, with Melissa Thompson, and the civic consequences of felon disenfranchisement, with Jeff Manza

HOWARD S. ERLANGER is professor of sociology, Voss-Bascom Professor of Law, and director of the Institute for Legal Studies at the University of Wisconsin, Madison. His current sociolegal work deals with lawyers' careers and with the implementation of law in the organizational context. Since 1982, he has been review section editor of *Law and Social Inquiry*, published by the University of Chicago Press.

ERIN KELLY is a Ph.D. candidate in sociology at Princeton University. Her dissertation examines the history of corporate family policies, including family leaves, child care programs, and flexible work arrangements, in U.S. workplaces. Her research interests also include gender theory, social movements, and social policy.

FRANK DOBBIN is associate professor of sociology at Princeton University. He is working on a book about how civil rights law has revolutionized employment practices. He is also examining, with Timothy Dowd, how

early industrial regulation shaped the emergence of modern business strategy.

DUNCAN J. WATTS received a Ph.D. in theoretical and applied mechanics from Cornell University in 1997 and is currently a postdoctoral fellow at the Santa Fe Institute. His primary research concerns the structure and dynamics of complex, interactive networks.

## Information for Contributors

**Editorial Procedures:** All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. Manuscripts are accepted for publication subject to the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal.* No paper known to be under consideration by any other journal will be reviewed by the *AJS.* In addition, the *AJS* is not a suitable vehicle for materials published elsewhere, in whole or in part. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his or her article without charge in any book of which he or she is the author or editor after it has appeared in the *Journal.*

**Manuscript Acceptance Policy:** While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the *Journal* has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the *Journal* in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

## Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, tables, and references—*double spaced,* allowing generous margins at the top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Do not use proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, beginning with the acknowledgment note as number 1. *Provide a word count on the cover page of the manuscript.*

2. Begin each table on a separate page. Refer to each table in numerical order in the text. Tables should not contain vertical lines or more than 20 two-digit columns or the equivalent.

3. Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.

4. Original or glossy prints of professionally drawn figures may be required if manuscript is accepted.

5. Include a brief abstract (not more than 100 words) summarizing the findings.

6. Submit *five copies* of the manuscript and abstract on white paper. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose a self-addressed stamped postcard to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article submitted to the *AJS* must be accompanied by a check or money order for $15.00, payable to the University of Chicago Press in U.S. currency or

its equivalent. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status. Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

**Format of References in Text:** All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61–65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of the Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (*a, b*) attached to the year of publication: (1965*a*).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons: (Duncan 1959; Gouldner 1963; Lipset 1964, pp. 61–65).

6. Format for references to machine-readable data files is

Davis, James Allan. 1978. GSS (General Social Survey). NORC ed. Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

**Format of References:** List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "REFERENCES." For examples, see text of articles in this issue.

(Rev. 3/98)

## CONTENTS

**Book Reviews**

## IN THIS ISSUE

ROBERT J. SAMPSON is Lucy Flower Professor of Sociology at the University of Chicago, senior research fellow at the American Bar Foundation, and scientific director for the Project on Human Development in Chicago Neighborhoods. The article in this issue joins his research interests in urban sociology and criminology. He also studies the life course and published (with John Laub) *Crime in the Making: Pathways and Turning Points through Life* (Harvard University Press, 1993).

STEPHEN W. RAUDENBUSH is professor of education in the School of Education and a senior research scientist at the Survey Research Center, both at the University of Michigan. He specializes in educational statistics and research design.

RUUD KOOPMANS is a senior researcher in the Department of Public Sphere and Social Movements at the Wissenschaftszentrum Berlin für Sozialforschung (WZB). He has published widely on comparative European politics, social movements and political participation, and German politics and society. His current research focuses on the politics of immigration and citizenship, and the claims making of xenophobic, antiracist, and ethnic minority groups in several Western European countries.

PAUL STATHAM is a senior research fellow in European political communication in the Institute of Communications Studies at the University of Leeds (UK) and a senior researcher in the Department of Public Sphere and Social Movements at the Wissenschaftszentrum Berlin für Sozialforschung (WZB). He has published on politics, media, and social movements, referring to Britain and Italy. His current research focuses on political claims making in the field of migration and ethnic relations in Europe.

CAROL MUELLER is associate professor of sociology at the new West Campus of Arizona State University and received her Ph.D. from Cornell University. She has edited or coedited three anthologies and numerous papers on social movements. Her most recent work examines theoretical and methodological issues in the fall of the German Democratic Republic.

MARC W. STEINBERG is assistant professor at Smith College. His writing has focused on historical class formation and conflict, repertoires of collective action and discourse, and he has recently published *Fighting Words: Working-Class Formation, Collective Action and Discourse in Early Nineteenth-Century England* (Cornell University Press, 1999). He is currently studying the role of law in class conflict and gender formation in mid-Victorian England.

JEFFERY M. PAIGE is professor of sociology at the University of Michigan. His most recent book is *Coffee and Power: Revolution and the Rise of Democracy in Central America* (Harvard University Press, 1997).

PATRICIA H. THORNTON is assistant professor in the Department of Sociology at Duke University. Her main research interests are in institutional and organizational change at different levels of analysis, including the

influences of politics, culture, and history on decision making in organizations, corporate governance, and the structural dynamics of markets and organizational populations. Her article, "The Sociology of Entrepreneurship," will appear in the 1999 *Annual Review of Sociology*.

WILLIAM OCASIO is assistant professor of organization behavior at the J. L. Kellogg Graduate School of Management and of sociology (by courtesy) in the Weinberg College of Arts and Sciences at Northwestern University. His main research interests are in the areas of power and politics in organizations, organizational decision making, and organizational and institutional change. He is currently studying the evolution of institutional logics and conceptions of control in U.S. industry.

## Information for Contributors

**Editorial Procedures:** All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. Manuscripts are accepted for publication subject to the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. In addition, the *AJS* is not a suitable vehicle for materials published elsewhere, in whole or in part. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his or her article without charge in any book of which he or she is the author or editor after it has appeared in the *Journal*.

**Manuscript Acceptance Policy:** While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the *Journal* has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the *Journal* in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

### Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, tables, and references—*double spaced,* allowing generous margins at the top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Do not use proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, beginning with the acknowledgment note as number 1. *Provide a word count on the cover page of the manuscript.*

2. Begin each table on a separate page. Refer to each table in numerical order in the text. Tables should not contain vertical lines or more than 20 two-digit columns or the equivalent.

3. Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.

4. Original or glossy prints of professionally drawn figures may be required if manuscript is accepted.

5. Include a brief abstract (not more than 100 words) summarizing the findings.

6. Submit *five copies* of the manuscript and abstract on white paper. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose a self-addressed stamped postcard to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article submitted to the *AJS* must be accompanied by a check or money order for $15.00, payable to the University of Chicago Press in U.S. currency or

its equivalent. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status. Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

**Format of References in Text:** All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61–65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of the Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (*a, b*) attached to the year of publication: (1965*a*).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons: (Duncan 1959; Gouldner 1963; Lipset 1964, pp. 61–65).

6. Format for references to machine-readable data files is

Davis, James Allan. 1978. GSS (General Social Survey). NORC ed. Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

**Format of References:** List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "REFERENCES." For examples, see text of articles in this issue.

(Rev. 3/98)

# Migration Patterns and the Growth of High-Poverty Neighborhoods, 1970–1990[1]

Lincoln Quillian
*University of Wisconsin, Madison*

Using geocoded data from the Panel Study of Income Dynamics, this article examines why the number of high-poverty neighborhoods in American cities has increased since 1970. The main findings are (1) the migration of the nonpoor away from moderately poor neighborhoods has been a key process in forming new high-poverty neighborhoods, although in the early 1980s increasing poverty rates were also important; and (2) African-Americans have moved into predominately white neighborhoods at a pace sufficient to increase their numbers there, but neighborhoods with increasing black populations tend to lose white population rapidly. Implications for theories of poor neighborhoods are discussed.

William Wilson's book *The Truly Disadvantaged* (1987) first pointed out that, starting in the 1970s, areas of concentrated urban poverty increasingly took on a different character than they had earlier in the century. Like the ethnic ghettos that have long interested urban sociologists, dwellers in modern poor urban neighborhoods are almost all members of minority races or ethnicities. Unlike older ethnic ghettos, however, Wilson argues that the minority-populated urban neighborhoods of the 1970s and

1980s contained an especially high concentration of poor families. He hypothesizes that one cause of this trend is that middle-class blacks in the 1970s and 1980s increasingly relocated to predominately white suburbs, leaving behind neighborhoods composed largely of poor or near-poor families.

Wilson's work leads empirical researchers to examine data to confirm or deny these suspicions. Investigations by Jargowsky (1994, 1997) have confirmed some of Wilson's hypotheses, finding that the proportion of the urban population living in census tracts in which at least 40% of the population is poor increased from 3% of the urban population in 1970 to 4.5% in 1990 (Jargowsky 1997, p. 38). Tests of Wilson's hypotheses about why this has occurred have been contradictory, and there remains a considerable debate about why poor urban neighborhoods have expanded so sharply.

An increase in the number of extremely poor neighborhoods can be thought of as resulting from a combination of two proximate causes: change in the number of poor persons and change in the tendency for persons of like poverty status to live close to each other. I decompose flows of persons among neighborhood and poverty status categories over time to examine how each of these proximate causes has influenced the number of high-poverty neighborhoods. This procedure sheds light on several explanations of the increase in neighborhood poverty.

Along the way, I consider evidence relevant to debates about the role of racial segregation in explaining concentrated urban poverty. I argue that studies of the role of racial segregation in forming high-poverty neighborhoods have not always clearly separated evidence about change over time from evidence about cross-sectional variation and have not fully considered the dynamics of neighborhood change. Research has found that racial segregation in American cities remains very high, even for high-income black families (Denton and Massey 1988; Massey and Denton 1993). This has been interpreted as inconsistent with Wilson's claims that middle-class blacks are migrating into white neighborhoods. A central finding of this article is that, when considered as part of a dynamic metropolitan setting, these apparently contradictory findings can be reconciled. Middle-class blacks have been moving into white neighborhoods at rates high enough to increase their numbers there, but declining white populations in neighborhoods with substantial black populations have prevented a large increase in the share of blacks in white nonpoor neighborhoods.

PAST THEORY AND RESEARCH

Three explanations predominate in the literature on the causes of high-poverty neighborhoods. These are that high-poverty neighborhoods result

from black middle-class flight from mixed-income neighborhoods, that high-poverty neighborhoods are the result of racial segregation, and that high-poverty neighborhoods are the result of the poor job prospects of inner-city workers (Massey, Gross, and Shibuya 1994). After defining a high-poverty neighborhood, I review each of these explanations with regard to what it can tell us about why the number of poor neighborhoods has increased over time. Then I turn to my own data analysis of patterns of migration to provide clues about these three theories.

Most prior research has defined high-poverty neighborhoods by a fixed cutoff based on the percentage of persons living in families with income below the federal poverty line, usually 30% or 40% (Wilson 1987; Jargowsky and Bane 1991). Defining poor neighborhoods based on a fixed cutoff point makes theoretical sense if there are thresholds in neighborhood poverty rates beyond which neighborhoods become substantially less livable. Jargowsky and Bane (1991) toured a number of cities, comparing block census maps showing poverty percentages to their impressions based on visual appearances. Areas with more than 40% poverty rates contained more dilapidated housing stock and seemed significantly more distressed than neighborhoods with poverty rates of 20%–40%, which had a more working-class character.[2] Following their work, I consider an extremely poor neighborhood to be a census tract in which more than 40% of persons are in families with incomes below the official poverty line.[3]

## Black Middle-Class Flight

In *The Truly Disadvantaged,* Wilson (1987) argues that one of the key forces that led to the increase in the number of extremely poor neighborhoods was the movement of middle-class residents, especially middle-class African-American residents, from mixed-income neighborhoods to suburban, white neighborhoods. With the departure of middle-class blacks from inner-city, mixed-income neighborhoods, the population left behind was

[2] Areas that were more than 40% poor also corresponded well to areas local census officials considered "ghettos." This standard has also been adopted by the U.S. Census Bureau, which refers to areas where more than 40% of the population is poor as "extreme poverty areas" (U.S Bureau of the Census 1995).

[3] Census tracts are small population units of 2,500–8,000 residents (average about 4,000) that are designed to be homogeneous with respect to population characteristics, economic status, and living conditions. They are drawn in such a way as to correspond roughly to what is normally thought of as a small neighborhood by people familiar with the local geography. For a discussion of how census tract boundaries are drawn and the advantages and disadvantages of considering census tracts as neighborhoods, see White (1987, pp. 18–20, 286–300).

"a much higher concentration of the most disadvantaged segments of the black urban population" (1987, p. 49). Wilson supports this conclusion by citing data on predominately African-American neighborhoods in Chicago that shows that from 1970 to 1980 the number of middle-class black families declined, but the absolute number of poor families remained roughly the same.

Several lines of research have investigated aspects of Wilson's claim. Studies of changes in the population of census tracts with increasing poverty rates come to conclusions broadly consistent with middle-class black out-migration. Studies in changes in racial segregation and the relationship between racial segregation and socioeconomic status (SES), on the other hand, have come to conclusions that seem inconsistent with middle-class black out-migration. I discuss these lines of research below and discuss possible explanations of the differences between the findings of these studies.

First, several studies have examined population changes associated with neighborhoods that became poorer between decennial censuses, and these studies support a connection between depopulation and increases in poverty rates. Case studies of the increase in high-poverty neighborhoods find that census tracts that had their poverty rate increase to beyond 40% poor usually show population losses, mostly due to a shrinking number of nonpoor families (Jargowsky and Bane 1991; Jargowsky 1997). Greene (1991) shows that there is a strong connection between loss of population and increase in tract poverty rates. Gramlich, Laren, and Sealand (1992) find that rates of migration among tracts imply that poor urban areas are gradually becoming poorer, blacker, and less densely populated. This research does not definitely establish that middle-class black out-migration is a cause of increases in the number of poor neighborhoods, but it is consistent with the possible importance of this factor.

The black out-migration thesis, however, has been challenged in a number of articles on residential racial segregation by Douglas Massey, Nancy Denton, and their colleagues. Using census extracts from major cities, they conclude that blacks of high SES (measured by income, education, or occupation) are only slightly less segregated from whites than low SES blacks (Massey 1979; Denton and Massey 1988). Furthermore, they show that SES is not related to greater suburbanization among blacks, a pattern different from Asians and Hispanics (Massey and Denton 1987). These studies demonstrate that African-Americans live largely in separate communities than whites, even at high levels of SES. This leads Massey and Denton (1987, p. 823) to conclude that "if the black middle class has abandoned the black poor, it has not been by moving to Anglo neighborhoods, at least not on a significant scale. Most blacks continue to reside in pre-

dominately black neighborhoods, even in cities with relatively large and affluent black middle classes."

These results leave open the possibility that middle-class blacks fled from inner-city, mixed-income areas to predominately black upper-income or middle-income areas. If this is the case, then we should see an increase in income segregation among blacks in the 1970s and 1980s. Several researchers have investigated this hypothesis, and the majority of these studies conclude that income segregation among blacks has been increasing (Massey and Eggers 1993; Jargowsky 1996).[4] These results support Wilson's hypothesis if we alter it to note that middle-class blacks are moving to middle-class black neighborhoods rather than white neighborhoods.

Yet the insights of most of these studies have been limited by the fact that they rely on static snapshots of the population from decennial census results. A few studies use longitudinal data from the Panel Study of Income Dynamics (PSID) to address urban poverty issues, but they are not studies designed to look at why the number of high-poverty neighborhoods has increased over time (e.g., Gramlich et al. 1992; South and Crowder 1997).[5] The decline in the number of middle-class residents of neighborhoods that became poorer could be because middle-class blacks migrated out *or* because the out-migration of poor and middle-class blacks was combined with the movement of many middle-class residents into poverty. This would show up in decennial census results as an increase in poverty rates and a decline in the number of middle-class blacks; yet, there would be no stronger a tendency for middle-class persons to migrate out of poor black neighborhoods than for poor persons. Thus, studies using cross-sectional samples at two points in time are ambiguous about the sources of the growth in poor neighborhoods because both poverty status and residence can change over time (Hughes 1990; Jargowsky and Bane 1991). To separate these and in general to provide a more complete view of the processes causing changes in neighborhood poverty rates, we need longitudinal data on families and the neighborhoods in which they reside.

---

[4] Farley (1991) is the one study that dissents from this conclusion. His study is the smallest in terms of number of metropolitan statistical areas (MSAs), and his measure of income segregation suffers from confounding changes in the black income distribution with changes in the degree of sorting into neighborhoods among black families, as discussed by Jargowsky (1996).

[5] I discuss the one study that uses longitudinal data to address some of these questions at length below.

Importance and Limitations of Racial Segregation Explanations

Massey and Denton's (1993) criticisms of the black out-migration thesis
are part of their wider criticism of Wilson's views on urban poverty. Their
main point is that Wilson has neglected the key role of racial segregation in
creating extremely poor neighborhoods. Indeed, one only needs to examine
statistics on the racial disparities in the population of high-poverty neigh-
borhoods to become convinced that there is a connection between racial
segregation and extremely poor neighborhoods. In 1990, about 14% of the
black urban population and 9.5% of the Hispanic urban population lived
in extremely poor neighborhoods. By contrast, only about 1% of the non-
Hispanic white urban population lived in extremely poor urban neighbor-
hoods (Jargowsky 1997).

As Douglas Massey, Nancy Denton, and their colleagues have made
clear in a number of articles, the disproportionate burden of ghetto pov-
erty on nonwhites is the result of two factors: the high degree of residential
segregation of nonwhites from whites and racial disparities in poverty
rates (see, e.g., Massey 1990). Nonwhite neighborhoods tend to have high
poverty rates because of the high rate of poverty among nonwhites. In
the population as a whole, about 10% of non-Hispanic whites are poor,
compared to about 30% of blacks and 25% of Hispanics. This means that
in a city with complete residential segregation, and no income segregation,
blacks would experience about a 30% neighborhood poverty rate—al-
ready a fairly high level of neighborhood poverty contact. In this racially
segregated city, even a low level of income segregation among African-
Americans is sufficient to create some neighborhoods with 40% or more
of their population with incomes below the poverty line. Whites in the
perfectly segregated city, in contrast, would experience only a 10% neigh-
borhood poverty rate. Although high levels of income segregation could
still lead to white neighborhoods that are at least 40% poor, in modern
American cities, levels of income segregation simply do not occur at high
enough levels to create more than a handful of extremely poor white
neighborhoods.

While racial segregation is critical to understanding the existence of
ghetto poverty, it is less clear that racial residential segregation can ex-
plain the change in the number of poor neighborhoods *over time*. Change
in racial segregation cannot explain the growth in the number of high-
poverty neighborhoods, because racial segregation declined slightly in the
United States from 1970 to 1990, while the number of high-poverty neigh-
borhoods increased (Jakubs 1986; Farley and Frey 1994). Massey and
Eggers (1990) and Massey and Denton (1993) argue that the effects of
economic changes has especially strong impacts on black neighborhoods
because of racial segregation. Although the empirical evidence on this

point is debatable, their simulations and theoretical arguments are convincing.[6] The *timing of the increase* in black neighborhood poverty rates, however, cannot be explained by changes in racial segregation.

## Economic Changes and Poverty Rates

Wilson (1987) hypothesizes that a second cause of increases in the number of high-poverty neighborhoods is economic change that worsened the employment and earnings prospects of blacks in inner-city neighborhoods. Wilson's later work (1996) emphasizes these demand-side economic factors as the fundamental causes of increasing urban poverty. His analysis is based on two theses from prior work on urban poverty. The spatial mismatch hypothesis, initially proposed by John Kain (1968), argues that jobs have increasingly moved away from urban central cities to suburban areas, leading to higher unemployment rates among blacks who live in inner-city neighborhoods. The deindustrialization hypothesis, associated with the work of several authors (e.g., Harrison and Bluestone 1981; Kasarda 1990), argues that there has been a decline in the number of inner-city factory jobs. As a result, wages and employment rates of inner-city residents have fallen relative to the wages and employment of suburban dwellers.

Although these theories are simple, they have been difficult to test. Several studies have tried to establish whether or not distance from available jobs contributes to unemployment (see Holzer [1994] and Teitz and Chapple [1998] for reviews). These studies have been primarily cross-sectional, often trying to estimate the relationship between measures of job access and unemployment and the extent to which racial differences in spatial mismatch can explain racial differences in unemployment rates. They do not establish the extent to which spatial mismatch and deindustralization can explain change in the unemployment rate of inner-city areas.

Studies that have considered changes in employment trends over time (e.g., Kasarda 1990) have been hampered by their inability to distinguish increases in rates of unemployment or poverty from the migration out of the employed or the nonpoor. Jencks and Mayer (1990, p. 220) conclude the failure to deal with class-selective migration has been "probably the single most important reason why we have learned so little about this subject in the two decades since Kain first advanced the spatial mismatch

---

[6] Massey and colleagues have argued that there is an interaction of changes in racial segregation and increasing poverty rates that has led to the increase in ghetto poverty (see, e.g., Massey and Eggers 1990). A reanalysis of Massey and Eggers's (1990) empirical evidence for this claim (Korenman, Sjaastad, and Jargowsky 1995), however, finds little evidence to support the presence of this interaction.

hypothesis." Solving this problem demands longitudinal data that can separate changes in employment or earnings resulting from demand-side job changes from changes in employment or earnings resulting from class-selective migration.

## Massey, Gross, and Shibuya's "Migration, Segregation, and the Concentration of Poverty"

One previous study, Massey et al.'s "Migration, Segregation, and the Concentration of Poverty" (1994), examines forces contributing to the growth of high-poverty neighborhoods, and Wilson's middle-class black out-migration thesis, using longitudinal data. I carefully consider these results here because my own work follows that of Massey, Gross, and Shibuya closely. Their article concludes that racial segregation is the central cause of ghetto poverty and that middle-class out-migration is at most a minor contributing factor. On this basis, they claim that Wilson was incorrect in asserting that black middle-class out-migration has significantly increased the concentration of poverty.[7]

Massey et al.'s (1994) investigation starts with an examination of transition probabilities among neighborhood types. The transition probabilities Massey et al. compute are the probabilities that a respondent will move to another neighborhood type conditional on their tract type of origin and destination and their race and poverty status. These probabilities indicate that poor blacks are moving out of poor neighborhoods at higher rates than nonpoor blacks, which, they argue, contradicts the black middle-class out-migration thesis.[8]

The results shown in Massey et al. (1994) are instructive in telling us about racial differences in patterns of movement among neighborhoods

[7] In discussing Wilson's thesis, they do not separate black migration into white neighborhoods from black migration into predominately black neighborhoods. Instead, they consider the importance of black migration into nonpoor neighborhoods overall.

[8] A problem with this result as a test of Wilson's middle-class out-migration thesis, however, is their definition of a movement "out." Massey et al. (1994) count any movement from a poor tract to any other tract as a movement out. Table 3 in Massey et al (1994) shows that most of these moves are from one poor black tract to another poor black tract. As they point out, poor blacks move out more often than nonpoor blacks largely because they tend to be renters and therefore tend to move more often. But this is not a test faithful to Wilson's intent: When Wilson refers to black "out" migration he is almost surely referring to movements out of the ghetto and into nonpoor areas, not any move originating in a poor census tract. A more reasonable way to adjudicate Wilson's hypothesis would be to compare a nonpoor and a poor black respondent on the probability that, if they live in a poor neighborhood at time $t$, they will reside in a nonpoor tract at time $t + 1$ (Jargowsky 1997).

TABLE 1

TRANSITION MATRIX FOR A HYPOTHETICAL
POPULATION DISTRIBUTED BETWEEN TWO TRACT
TYPES

| ORIGIN NEIGHBORHOOD TYPE | DESTINATION NEIGHBORHOOD TYPE | |
|---|---|---|
| | Nonpoor | Poor |
| Nonpoor . .. .. ..... . ... . . .. | .45 | 55 |
| Poor  .. ......... .. ... ............ ... . | .3 | .7 |

NOTE —Numbers represent the probability of being in the column tract type at time $t + 1$ given residence in row tract type at time $t$

of different types. Yet they have a sharp limitation in that these rates alone do not tell us about change in the population over time. Massey et al.'s analysis tests Wilson's hypothesis only as it applies to cross-sectional changes, not to explaining change in the number of poor neighborhoods over time.

To understand why this is true, consider the hypothetical population in table 1. This table shows a matrix of transition rates among two neighborhood types. The numbers represent the probabilities of moving to the column neighborhood type given the row origin neighborhood type. In this example, there are only two neighborhood types: poor and nonpoor.[9] Table 2 shows the proportion of the population living in each neighborhood type given at a hypothetical starting point and then after one application of the transition matrix, three applications of the transition matrix, and, finally, the stable distribution that the population will approach given repeated application of the transition matrix.[10] In the example, although the probability of moving into a poor neighborhood is greater than the probability of exiting it for both residents of poor and nonpoor neighborhoods, the proportion of the population residing in the poor neighborhood type is *declining* over time. Because most of the population lives in poor neighborhoods to begin with, even a low probability of moving to a nonpoor area is compatible with an increase in the proportion of the population in nonpoor neighborhoods.

Massey et al. (1994) find that the nonpoor blacks are less likely to enter

[9] These results generalize to $N$ neighborhood types.

[10] For discussions of rates and the stable population they imply see Rogers (1968), Keyfitz (1977), or Boudon (1973).

TABLE 2

CHANGE IN POPULATION DISTRIBUTION OVER TIME AS TRANSITION
MATRIX RATES ARE APPLIED

|  | NEIGHBORHOOD TYPES | |
| --- | --- | --- |
|  | Nonpoor (%) | Poor (%) |
| Starting population distribution (time $t$) ............. | 20.0 | 80.0 |
| Population distribution at time $t + 1$ .................. | 33 0 | 67.0 |
| Population distribution at time $t + 3$ ............ ..... | 35.2 | 64.8 |
| Stable population distribution ... ...... ..... .... ........ | 35.3 | 64.7 |

nonpoor areas than poor areas. The implication of the example in tables 1 and 2 is that their result is consistent with the possibility that the number of blacks in white neighborhoods is increasing over time. If there are sufficiently few blacks in white neighborhoods to begin with, then even a low probability of black migration to a nonpoor white neighborhood can be consistent with an increase in the size of the black nonpoor population in white neighborhoods over time. More generally, because population change is a complex function of both the rates of change and the population distribution at a given point in time, we can say nothing definitive about how a population is changing over time just by looking at transition rates. Nor can we conclude that because group A has a lower rate of entry to a neighborhood than group B, group A is declining in size in that neighborhood relative to group B. A better way to examine the implications of these rates for the distribution of people across neighborhood types is to examine flows, a method I use below.[11]

The other technique Massey et al. (1994) rely on is a simulation. They begin with a hypothetical city composed of neighborhoods of economic and racial composition similar to the city of Chicago. They then "age" the population five years by applying transition probabilities, estimated from the PSID, five times. They show that giving blacks and whites equal destination probabilities has a much larger impact on the average poverty rate experienced by blacks than eliminating middle-class black out-migration

[11] Preston and Campbell (1993) made a related point in discussing how differences in rates of fertility among people of different IQ classes will result in changes in the distribution of IQ over time Another way to examine change over time based on rates is to compute the stable population distribution implied by the rates and then compare this to the observed population distribution.

from poor areas or eliminating socioeconomic mobility. On this basis, they conclude that racial segregation is ultimately more important than black out-migration in explaining the concentration of black poverty.

There are two caveats to the results from the simulation. First, although the procedure of assigning black probabilities to whites is an interesting hypothetical, it, again, does not address historical trends over time. Massey et al. (1994) do not examine what actual rates of movement from white to black tracts or vice-versa imply about change over time in the distribution of blacks in poor neighborhoods. A second limitation is that they begin their simulation with a hypothetical city with a racial and economic distribution of neighborhoods similar to Chicago. Since the degree of change toward the equilibrium distribution is related to the difference between the starting and stable distribution, choosing a city with a different degree of segregation as the starting point would lead to different results. In any event, it is difficult to know if the results would be different enough to change their conclusions without reanalyzing the data with different starting distributions.

To summarize, the results Massey et al. (1994) present are persuasive for illustrating the existence of great racial differences in patterns of movement between neighborhood types and for illustrating the crucial importance of racial segregation for maintaining extremely poor urban areas. Like other studies of transition rates, however, their results do not provide much insight into questions about changes in racial segregation over time or the possible sources of the increase in high-poverty neighborhoods in the 1970s and 1980s. In what follows, I perform a number of different analyses to determine what the PSID data implies about the growth in high-poverty neighborhoods in the 1970s and 1980s.

## Cross-Sectional and Temporal Explanations of Urban Poverty

The two explanations of high-poverty neighborhoods most directly relevant for looking at change over time are black middle-class flight (Wilson 1987) and the worsening economic circumstances of inner-city workers (Wilson 1987, 1996). Racial segregation (Massey and Denton 1993) is an important background condition that is a prerequisite for the existence of high-poverty neighborhoods, but in itself it is not an explanation of why there has been a growth in the number of high-poverty neighborhoods. The strengths of these explanations reflect the tendency of Wilson and Massey to address slightly different dependent variables: Wilson's work focuses on explaining change in the number of poor neighborhoods over time, while Massey and his colleagues' work focuses more on explanations of the existence of ghetto poverty overall, in most cases using statistical

methods that explain cross-sectional variation. In light of their slightly different emphases, it is not surprising that their opposing theses have not yet resulted in an enlightening synthesis.

For reasons I have spelled out above, empirical work leaves many unanswered questions about how these factors have influenced the growth in high-poverty neighborhoods. Most empirical research has either relied on cross-sectional snapshots, and suffers some attendant limitations, or has focused on explaining cross-sectional variation rather than change over time. Below, I analyze longitudinal data from the PSID to examine what it tells us about changes over time in the number of poor neighborhoods.

## METHODS

My basic strategy is to decompose changes in the number of nonpoor whites, poor whites, nonpoor blacks, and poor blacks in the PSID living in poor neighborhoods into three sources of change: movement, neighborhood change around respondents who do not move, and change in poverty status among stayers. I also separately consider respondents who both changed poverty status and switched neighborhood type.

### Data

To investigate the causes of the growth in concentrated urban poverty, I rely on data from the PSID. The PSID has followed approximately 5,000 families and their descendants with yearly interviews since 1968. To study . changes in neighborhoods over time, I have matched data on PSID respondents to data on census tract characteristics from the 1970, 1980, and 1990 censuses.[12] The PSID sample originally included an oversample of poor families; I employ the PSID sampling weights for all analyses in this article to make the results representative of the U.S. population.

Unfortunately, the PSID address tapes for 1969, 1975, 1977, and 1978 were missing when my data extract was compiled. Although geocodes for 1975, 1977, and 1978 are now available, they are only available using 1990 census geography. Since 1990 is 12–15 years later than 1975–78, using this data to represent tract characteristics for 1975–78 would probably be highly unreliable, and instead I exclude the years 1975–78. As a result, my PSID extract uses data from 1970–74 and 1979–90.

The PSID data and census geocodes are available for respondents at single-year intervals. Data on the neighborhoods in which these respon-

---

[12] See n. 2 above for more information on census tracts.

dents live, however, is only available from the census at 10-year intervals. This leaves the problem of how to impute neighborhood characteristics for years between censuses. The imputation procedure is discussed at greater length in the next section.

Since the PSID sample before 1990 has very few Latinos or Asians, I examine black and white PSID sample members only. I divided the respondents into four groups: white nonpoor, white poor, black nonpoor, black poor. I consider a PSID respondent to be poor if they are a member of a family whose three-year average posttransfer income is less than 125% of the federal government poverty needs standard.[13] I use a three-year moving average of income to needs because much of the transitory fluctuation in income is measurement error. In addition, many families whose income is below the poverty line in only one year are not "poor" in any meaningful sense since they have sufficient assets and social support to avoid hardship during their period of low income (Rodgers and Rodgers 1993).

To analyze the relation between individual and neighborhood characteristics, I create three income categories: nonpoor (less than 20% of population in the tract is in households below the federal poverty threshold), moderately poor (20%–39.9% poor), and extremely poor (40%+ poor). I also create three racial tract types: white (less than 30% of population black), mixed (less than 70% black), and black (70%–100% black). Cross-categorizing the neighborhood poverty and racial types forms nine cells. In addition, I add a tenth cell for nonmetropolitan residence and an eleventh cell for individuals who live in metropolitan areas but do not have tract addresses because they live in a nontracted metropolitan area or because they provided an address that the PSID was unable to assign to a single tract.

The number of respondents in some of these 11 neighborhood types, however, was too small to support an analysis. As a result, I further collapsed these 11 categories down to eight categories. In so doing, I collapsed together white moderately poor and extremely poor neighborhoods (there are very few white extremely poor neighborhoods). I also collapsed all poverty levels of racially mixed neighborhoods into one category, racially mixed (30%–70% black). Table 3 shows the weighted percentage of person years for each neighborhood type separately for the four race by poverty status categories using the full pooled sample (1970–74, 1979–90). The unweighted number of person years is also shown in table 3.

[13] Since there is less underreporting of income in the PSID than in the census, using the 125% poverty line poverty rates in the PSID sample are comparable to poverty rates in the census or CPS using the 100% poverty line (Hill 1992)

## TABLE 3

PERCENTAGE AND NUMBER OF PERSON YEARS BY NEIGHBORHOOD TYPE, POOLED PSID 1970–74 AND 1979–90 DATA

| NEIGHBORHOOD TYPE | WHITE NONPOOR | | WHITE POOR | | BLACK NONPOOR | | BLACK POOR | |
|---|---|---|---|---|---|---|---|---|
| | % of Person Years | Person Years | % of Person Years | Person Years | % of Person Years | Person Years | % of Person Years | Person Years |
| White nonpoor* | 47.2 | 45,964 | 21.9 | 2,166 | 11.3 | 4,919 | 5.0 | 1,814 |
| White moderately poor and white extremely poor† | 3.0 | 3,126 | 10.1 | 1,039 | 3.2 | 1,059 | 2.1 | 668 |
| Racially mixed‡ | 1.5 | 1,896 | 2.8 | 403 | 19.8 | 7,655 | 15.2 | 4,077 |
| Black nonpoor§ | .1 | 99 | .1 | 17 | 10.4 | 5,805 | 4.2 | 1,886 |
| Black moderately poor‖ | .2 | 300 | .5 | 153 | 20.4 | 12,142 | 17.6 | 8,384 |
| Black extremely poor¶ | .0 | 107 | .0 | 12 | 6.9 | 4,593 | 13.9 | 7,190 |
| Nonmetropolitan | 36.7 | 38,310 | 53.4 | 5,136 | 20.3 | 8,727 | 33.9 | 8,047 |
| Nontract metropolitan | 11.3 | 12,450 | 11.1 | 1,322 | 7.6 | 4,502 | 8.2 | 3,498 |
| Total | 100.0 | 102,252 | 100.0 | 10,248 | 100.0 | 49,402 | 100.0 | 35,564 |

NOTE.—Numbers are based on imputing tract characteristics for intercensal years based on tract-to-tract matching; see methods section "% of person years" are given as weighted and "person years" as unweighted

\* 0%–30% black, 0%–20% poor.
† 0%–30% black; 20%–100% poor
‡ 30%–70% black, 0%–100% poor
§ 0%–30% black; 0%–20% poor.
‖ 70%–100% black, 20%–40% poor
¶ 70%–100% black, 40%–100% poor.

Imputation of Intercensal Years

One of the goals of this article is to measure the impact of neighborhood changes around PSID respondents on movement among census tract types. This requires some care because many census tracts' boundaries change between censuses, and the method of interpolation for these tracts may influence results. To deal with this problem, I tried three different methods of imputing census tract racial makeup and poverty makeup for intercensal years.

In the first method, I assign neighborhood characteristics to PSID respondents based on the nearest census year. Tract characteristics for 1971–74 are assigned based on 1970 tract data, tract characteristics for 1979–84 are assigned based on 1980 tract data, and characteristics for 1985–89 are assigned based on 1990 census tract data. This procedure treats tract characteristics as if they are fixed at particular census years. I call this the "nearest-census-year" method.

In the second method, I matched all census tracts that did not change or that only had minor changes in their boundaries from 1970 to 1980 using the Census Bureau's 1970–80 census tract match file (U.S. Bureau of the Census 1983a). I similarly matched census tracts in 1980 with the corresponding tracts in 1990 using the Census Bureau's 1980–90 tract match file (U.S. Bureau of the Census 1992a).[14] Then I filled in values for years between censuses for the matched tracts using linear interpolation. For years that a respondent is a resident of the approximately 15% of tracts that had more than minor boundary changes between censuses, tract data is missing. Person years in these tracts are excluded from the analysis. This method I call "tract-to-tract interpolation" because it is based on matching tracts that did not change significantly; it is the basic method used in the tables shown in the results section.

The potential problem with this method is possible selection bias in excluding tracts that have had boundary changes. Given the Census Bureau's rules for drawing tract boundaries, one would guess that census tracts that had boundary changes were also the tracts that had large changes in the demographic makeup of their neighborhood populations.[15] Excluding these tracts might tend to systematically underestimate the ef-

---

[14] "Minor changes" are defined following the Census Bureau's definition in the tract matching files. For 1970 and 1980, minor changes are those that involved a gross population shift of less than 100 persons (U.S. Bureau of the Census 1983b). For 1980 and 1990 minor changes are those that affected less than 2.5% of the 1990 population of the tract (U.S. Bureau of the Census 1992b).

[15] The rules and procedures for drawing tract boundaries are summarized in White (1987, app. B).

fect of neighborhood changes because it eliminates many of the neighborhoods that had the most radical demographic shifts. To deal with this problem, I compared results using the tract-to-tract method with results employing a third method less sensitive to this problem.

The third method takes advantage of the fact that the PSID provides both 1970 and 1980 geocodes for PSID respondents for the period 1970–85. I fill in intercensal years using linear interpolation between the census tract that the respondent's dwelling was in using the 1970 geocode and the census tract that the respondent's dwelling was located in using the 1980 geocode. If the respondent's census tract changed between censuses, then the 1970 and 1980 tracts will have different boundaries. If boundaries of tracts really represent neighborhoods, then changes in tract boundaries represent change in neighborhood boundaries, and this method of imputation probably makes sense. If local census tract committees are drawing tract boundaries more arbitrarily, on the other hand, then this procedure would probably overstate the extent of neighborhood change because many apparent neighborhood changes are really arbitrary shifts in boundaries. This procedure I call "place-to-place interpolation," because it is based on interpolation between the tracts that include addresses in different censuses years (which in some cases are different tracts).

I have computed the basic results for this article using all three methods. The tables show the results using matching based on the tract-to-tract method of imputation. The disadvantage of place-to-place interpolation is that the geocodes needed to do it are not available for the years 1985–90. Tables showing flows using the tract-to-tract and the place-to-place method, broken down in the three basic periods (1970–74, 1979–84, and 1985–90), are available from the author upon request.

The results using the place-to-place method are generally very similar to results using the tract-to-tract method and in no way alter the substantive results of this article.[16] Estimated flows due to neighborhood change tend to be slightly larger using the place-to-place method than the tract-to-tract method.

A Method to Study Change over Time: Flows

In the literature review above, I argued that simply examining transition rates and their correlates is not a method well suited to tell us about

---

[16] The nearest-census-year method estimates sometimes differ by a larger amount because they do not account for changes in neighborhoods over time. This is what we would expect if neighborhoods are changing over time and suggests that conclusions of studies using the nearest-census-year method to study change over time may be misleading.

changes over time. There is no way to tell if the number of persons living in a particular neighborhood type is increasing or decreasing just by perusing rates. A better method is to consider flows. The flow is the probability of changing to another neighborhood (or making whatever other transition) times the size of the population that is at risk—that is, the flow is the size of the population moving in or out. Flows have the advantage that they are unambiguous about change over time. Positive net flows indicate the population in an area is increasing over time, negative net flows that it is decreasing over time.

Using the PSID, I examine the net flows of respondents based on their race and poverty status into and out of the eight neighborhood types. The net flow into a particular state is the share of the population that enters that state minus the share of the population that exits that state. To define it more precisely, first define a consecutive person year (CPY) as a sequence of two years in which we observe a PSID respondent and have valid data on neighborhood characteristics based on PSID geocoding. Then we can define $\text{in}_{nptr}$ as the flow into neighborhood type $n$ (the eight neighborhood types are listed in table 1) for persons of poverty status $p$ (nonpoor, poor) and racial group $r$ (black, white) in time period $t$ (1970–74, 1979–84, 1985–90). Define the flow as equal to

$$\text{in}_{nptr} = \text{entrances}_{nptr}/\text{CPY}_{tr},$$

where $\text{entrances}_{nptr}$ is the number of consecutive person years that move respondents who are members of neighborhood $n$ and poverty status $p$ out of that state (by exiting neighborhood type $n$, poverty status $p$, or both). $\text{CPY}_{tr}$ is the total number of consecutive person years observed for members of racial group $r$ during time period $t$. Thus, the flows I am using are normalized to be a percentage of all consecutive person years observed by members of the racial group during the given period. We can interpret $\text{in}_{nptr}$ as the average proportion per year of the $r$th racial group (blacks or whites) that entered neighborhood type $n$ and poverty status $p$ in period $t$. In a similar way, we can define movements out as

$$\text{out}_{nptr} = \text{exits}_{nptr}/\text{CPY}_{tr},$$

where $\text{exits}_{nptr}$ is the number of consecutive person years that move respondents who are in both neighborhood type $n$ and poverty status $p$ out of neighborhood type $n$ or poverty status $p$, and $\text{CPY}_{tr}$ is the total number of consecutive person years observed for members of racial group $r$ during time period $t$. We can interpret $\text{out}_{nptr}$ as the average proportion per year of racial group $r$ (blacks or whites) that exited neighborhood type $n$ and poverty status $p$ in period $t$.

Define the net flow, $f$, as the difference between these two quantities:

$$f_{nptr} = \text{in}_{nptr} - \text{out}_{nptr},$$

where $f$ is the net flow into (or out of, if the number is negative) the $n$th neighborhood type for persons in the $p$th poverty status for persons of race $r$ during time period $t$. The net flow is the number presented in the tables below. It can be interpreted as the average percentage per year of the racial group that moves in or out of the race by poverty status state in the specified period; a positive percentage indicates net movement in, a negative percentage a net movement out. For many PSID respondents, more than one consecutive person year is counted in computing a total for a given time period—if two consecutive person years are observed for a given individual during a period, then each of these consecutive person years is counted separately in computing the flow totals. This does not bias these estimates of population change, but it does create special problems in computing accurate standard errors of the flows, a problem I deal with at greater length in the next section.

An example might help illustrate the calculation. If 2% of the consecutive white person years in the PSID data are moves into a white nonpoor neighborhood by a white nonpoor person, and 1.5% of the consecutive white person years in the PSID are moves out of a white nonpoor neighborhood by a white nonpoor person, then the net flow into the neighborhood type by white nonpoor persons would be .5%.

I calculate flows separately for residents of each of the six basic neighborhood types, and I calculate net flows as percentages of the racial group separately for nonpoor blacks, poor blacks, nonpoor whites, and poor whites.

We can further decompose the flows based on entries or exits for particular reasons. I separately consider four ways in which PSID respondents can switch among neighborhood types. First, they can simply move from one neighborhood type to another neighborhood type. Second, they can stay still while the neighborhood type changes around them. Neighborhoods can become poorer or blacker (or less poor or more white) while the respondent does not move. Third, respondents can change poverty status while staying in their current neighborhood type. That is, they can become poor or become nonpoor, thus entering or exiting these subpopulations. Finally, respondents can both change poverty status (enter or exit poverty) and change their neighborhood because of moving or neighborhood change. When both of these events occur between PSID interviews, I cannot distinguish which occurred first; accordingly, I consider these events simultaneous and put people experiencing simultaneous events into their own category. I break down the flows into these four component sources of change. The total net flow of a neighborhood type is equal to the sum of its net flows due to each of these four reasons.

Several other sources of population change can also influence the size of populations in poor neighborhoods but are not accounted for in these

models. For example, births and deaths are not accounted for. Entries into institutions also are not included because the PSID does not gather data on people in institutions.[17] Flows of people who enter or exit neighborhoods for these reasons are not counted as consecutive person years. Changes in the population composition of neighborhoods because of these processes remain potential areas to investigate in future research.

Inferential Statistics

Using formulas that assume simple random sampling to compute standard errors is clearly inappropriate for computing standard errors for the flows calculated in this article. There are a number of reasons why this is the case. First, the PSID sample was initially drawn using a stratified and clustered design and so was not a simple random sample to begin with (Hill 1992). Second, the transitions that are used to compute flows for each period often include multiple transitions for a single individual, violating the assumption that the observations are independent. Third, by definition, individuals within the same family unit reside at the same location and share the same poverty status. There is perfect clustering within families on these attributes. Standard error calculations using simple random-sample formulas with these data are likely to be substantial underestimates (Wolter 1985).

To deal with this problem, I used variables identifying stratum and sampling-error computation units in the PSID to compute corrected standard errors using the Taylor series linearization (or delta) method (Kalton 1979; Wolter 1985). The estimates do not make any assumptions about the structure of the errors within the clusters.[18] Since all the person years of a particular individual and of their family members are within the same cluster, this method effectively accounts for the clustering caused by multiple observations on individuals and family members. The standard er-

[17] Institutions include armed forces barracks or quarters, college dorms, hospitals, prisons, and residential communities for members of religious orders. Between 1.5% and 3% of the PSID sample in any given year is in one of these institutions (Survey Research Center 1998).

[18] This was done using the survey data commands in the statistical package Stata 5.0. Because these estimates make no assumption about correlations among observations within the first-stage probability sampling units (and thus do not use information about randomness introduced into the design by sampling at later stages), they probably are slightly upwardly biased. In a sense, these estimators take the approach that the only randomness in the sample is introduced by stratum/probability sampling unit selection (see StataCorp 1997, sec. 36 2.1). The standard errors do not, however, account explicitly for the interpolation procedure used to fill in intervening census tract years.

TABLE 4

AVERAGE NET POPULATION FLOWS AMONG NEIGHBORHOOD TYPES DUE TO MOVEMENT
AND NEIGHBORHOOD CHANGE POOLED DATA FROM 1970–74 AND 1979–90

| | MOVEMENT | | NEIGHBORHOOD CHANGE | |
|---|---|---|---|---|
| NEIGHBORHOOD TYPE | Nonpoor | Poor | Nonpoor | Poor |
| Blacks. | | | | |
| White nonpoor ..................... ...... .. .............. | .337* | .018 | −.275* | −.072 |
| | (.0519) | ( 0646) | (.0748) | ( 0404) |
| Racially mixed ................... ........................ | .094 | .122 | .102 | 110 |
| | (.1525) | (.0670) | (.1038) | (.1599) |
| Black nonpoor ... .................. ... .. .. . . .. | −.110 | .055 | −.265* | .011 |
| | (.0736) | (.0341) | (.1300) | (.0556) |
| Black moderately poor .............. .. ...... . ... | − 147 | −.093 | .253 | −.173* |
| | (.0979) | (.0925) | (.1821) | (.0644) |
| Black extremely poor ......... . .... . ............... | −.089 | −.046 | .176* | 239* |
| | (.0490) ·, | (.1071) | (.0604) | (.0913) |
| Whites: | | | | |
| White nonpoor .. .. ...................... ....... ......... | .028 | 019 | −.191* | −.014* |
| | (.0392) | (.0126) | (.0517) | ( 0057) |
| White moderately and extremely poor ...... | −.062* | .004 | .096* | .005 |
| | (.0233) | (.0133) | ( 0184) | ( 0057) |
| Racially mixed ..... .. . .. .......................... ... | −.033 | −.002 | .087* | .004 |
| | (.0418) | ( 0049) | (.0318) | (.0044) |

NOTE —The tables includes only consecutive person years in which poverty status for the respondent did not change. The figures are given in percentages of the total black population or white population, respectively. So a .5 in the movement section of the top panel, for instance, represents 5% more of the black population moving into the neighborhood type than moving out. SEs are in parentheses
\* = flow/SE > 2

rors computed this way were on average about 2–2.5 times larger than standard errors computed under the assumption of simple random sampling.

RESULTS

Tables 4 and 6 contain estimates from the PSID sample of the net population flows resulting from movement, neighborhood change, and poverty status change. The results are shown separately for four groups: black nonpoor, black poor, white nonpoor, and white poor. Estimates were computed for each of the eight neighborhood types shown in table 3. However, I do not show results for neighborhood types that contained too few cases to permit an analysis (African-Americans in white poor neighborhoods, whites in black neighborhoods) or that contained no results of relevance

(nonmetropolitan areas, nontract metropolitan areas). The PSID sample size is not extremely large and is clustered; thus, many of the sample estimates are somewhat imprecise. Unless a net flow is substantial, it will not show up as statistically significant in the tables.

Table 4 shows estimates of average net flows of whites and black PSID respondents among neighborhood types resulting both from movement and neighborhood change. The flows due to movement are shown in the two left-most numeric columns. The numbers are given as percentage of the black or white population, respectively. For instance, in the top half of table 4, the .337 in the white nonpoor row indicates that on average .337% more of the black nonpoor population moved into that neighborhood type in each year than moved out. A positive number in table 4 indicates that the flow into the neighborhood is increasing the size of the group over time (more entrances than exits), while a negative flow indicates that the flow contributes to a decline in size over time (more exits than entrances).

Table 4 shows the movement and neighborhood change flows pooled across years because the flows usually did not differ by time period. Instances where the flows did differ by time period are discussed in the text. The full tables showing all flows broken down separately by time period (1970–74, 1979–84, and 1985–90) are available from the author upon request.

Movement

The results in the two left-most columns of table 4 show the net flows resulting from geographical movement to a different neighborhood type. This allows me to assess which neighborhood types have had more PSID respondents move in than move out and vice-versa. Only moves that take a respondent to a different neighborhood type are counted as movement for purposes of this statistic; relocation to another dwelling within the same neighborhood type is not counted as moving.

The largest population flow is the movement of nonpoor blacks into white nonpoor neighborhoods. Substantially more nonpoor blacks move into white nonpoor neighborhoods each year than move out. From table 4, on average about .337% more of the African-American population moves into white nonpoor neighborhoods each year than moves out. Although this might not seem like a large flow at first glance, over a 10-year period if the flow remains constant this represents a movement of about 3.3% of the black population into white nonpoor neighborhoods. Since about 11% of the black nonpoor population lives in white neighborhoods in the PSID sample, this would represent about a 32% increase in the share of nonpoor blacks living in white neighborhoods over 10 years.

KP-787

TABLE 5

MOVEMENT OF AFRICAN-AMERICANS INTO AND OUT OF WHITE NONPOOR
NEIGHBORHOODS, AVERAGE NET FLOWS AND PROBABILITIES BY PERIOD

| PERIOD | FLOW INTO OR OUT OF WHITE NONPOOR NEIGHBORHOODS | | PROBABILITIES OF MOVING INTO A WHITE NONPOOR NEIGHBORHOOD | | PROBABILITIES OF MOVING OUT OF A WHITE NONPOOR NEIGHBORHOOD | |
|---|---|---|---|---|---|---|
| | Nonpoor Blacks | Poor Blacks | Nonpoor Blacks | Poor Blacks | Nonpoor Blacks | Poor Blacks |
| 1970–74 .......... ... | −.054 | −.089 | 012* | 004* | .095* | .146 |
| | (.1556) | (.1112) | ( 0018) | (.0013) | (.0303) | ( 0931)' |
| 1979–84 ...... .. . | .407* | .007 | 031* | .012* | .094* | .208* |
| | (.1148) | (.0302) | (.0062) | ( 0046) | (.0149) | ( 0368) |
| 1985–90 .... . ....... | .503* | .088 | 036* | .010* | .094* | .128* |
| | (.0647) | (.0903) | (.0071) | (.0011) | (.0135) | (.0142) |

NOTE —Flows are as a percentage of the total black population, see notes to table 4 for an explanation
of population figures The entrance columns give the probability that a black respondent not residing in
a white nonpoor neighborhood will enter a white nonpoor neighborhood during the next year from any
other neighborhood type. The exit columns give the probabilities that a black respondent in a white
nonpoor neighborhood will exit during the next year. SEs are in parentheses.
 * = coefficient/SE > 2

These results support Wilson's (1987) contention that nonpoor blacks are
moving out of black metropolitan areas into predominately white neigh-
borhoods at a pace sufficient to increase their numbers there.

Flows for poor African-Americans are shown in the top half of table 4.
The net flow shown in the movement row is close to zero (.018) and not
statistically significant. Unlike for nonpoor blacks, poor blacks are moving
into white nonpoor neighborhoods about as often as they are moving out.
These results support Wilson's contention that nonpoor blacks are becom-
ing more spatially separated from poor blacks because nonpoor blacks
are relocating to white areas faster than poor blacks.

Table 5 provides more detail on the movement of African-Americans
into and out of white nonpoor neighborhoods. The table shows a break-
down of the flow for white nonpoor neighborhoods by time period and
probabilities of entry and exit. There is a net positive flow into white
nonpoor neighborhoods for 1979–84 and 1985–90 but not for 1970–74.
This suggests that the flow of African-Americans into white neighbor-
hoods increased during the late 1970s, but because the 1970–74 flow esti-
mate is imprecise and fragile to the method of imputation, this conclusion
cannot be made with high confidence.[19]

[19] The estimated flow of nonpoor blacks into white neighborhoods for 1970–74 is one
of the only flows that changed notably depending on the method of imputation of tract

The probabilities of entry and exit given in table 5 show that the probability a black nonpoor family will enter a white nonpoor neighborhood is about three times greater in the later two periods (1979–84 and 1985–90) than the earlier period (1970–74).[20] The results also show nonpoor blacks are substantially more likely to move into white nonpoor neighborhoods than poor blacks in all three time periods. Again, these results are consistent with Wilson's hypothesis of increasing spatial separation of poor and nonpoor blacks. More tentatively, the results suggest that the rate of movement of blacks into white neighborhoods increased in pace during the late 1970s.

If nonpoor blacks have net positive movement into white nonpoor neighborhoods, what neighborhood types are they moving from? The net positive flow because of movement of nonpoor blacks into white nonpoor neighborhoods originates from several neighborhood types. The point estimates in table 4 show negative flows from black nonpoor, moderately poor, and extremely poor tracts. Although the individual coefficients in table 4 are not statistically significant at conventional levels, the sum of the flows out of black poor neighborhoods is statistically significant ($P <$ .05).

The movement patterns for whites are also shown in table 4. Like for African-Americans, the results suggest that whites are moving out of poor neighborhoods and into nonpoor neighborhoods. The size of the flows are not nearly as large for whites as for African-Americans, though, and some flows are not statistically significant.

Overall, there is clear evidence that both nonpoor black and white PSID respondents have tended to move out of poor neighborhoods and into nonpoor white neighborhoods. By far, the most substantial flow is of nonpoor blacks, who are moving into white nonpoor neighborhoods at a rate sufficient to substantially increase their numbers there in a 10-year period. The most likely interpretation of these flows is that nonpoor blacks are moving in patterns to avoid neighborhoods with moderate to high rates of poverty, a path that often takes them into white neighborhoods. This migration is probably intended to avoid problems associated with high rates of neighborhood poverty such as high rates of crime, poor physical upkeep of the neighborhood housing stock, and inadequately funded gov-

characteristics. Flows based on the nearest-census-year method show a movement of nonpoor blacks into white nonpoor neighborhoods of .479% per year for 1970–74 ($P <$ .05); flows using the tract-to-tract method show a flow in of .233 (not statistically significant). The conclusion that the increase in the probability of entry into white neighborhoods occurred between 1974 and 1979 should be regarded with caution because of uncertainty about the 1970–74 flow.

[20] The caveats about the 1970–74 flow in n. 19 apply to the 1970–74 probabilities of moving in as well

ernment services (Krivo and Peterson 1996; Morenoff and Sampson 1997; Skogan 1990).

## Neighborhood Change

Just examining the movement of PSID respondents leaves out the other important way in which persons can switch census tracts. As I discussed in the methods and data section, individuals can change from one neighborhood type to another in two ways: by moving to a new neighborhood type or by staying put while the neighborhood type changes around them. Net flows among tract types due to neighborhood change around respondents are shown in table 4 under the heading "neighborhood change."

As is true for movement, the most substantial net flow because of neighborhood change in the table is for nonpoor blacks living in white nonpoor neighborhoods. In the "neighborhood change" row of the top half of table 4, .275% more of the nonpoor population is exiting white nonpoor neighborhoods each year than entering, because their neighborhoods become blacker around them. For the nonpoor African-American population, changes in neighborhoods are a cause of net movement out of white nonpoor census tracts. In fact, these flows out due to neighborhood change largely counteract the movement of African-Americans into white nonpoor neighborhoods.

Neighborhoods that switch out of the white nonpoor neighborhood category can do so either by becoming white poor neighborhoods or racially mixed neighborhoods. We can break down this total net flow out of white nonpoor neighborhoods of .275% into net neighborhood change with each of these two neighborhood types (appendix table A1 shows a complete breakdown). The net neighborhood change flow of African-Americans from white nonpoor to white poor neighborhoods due to neighborhood change is .082, while the net neighborhood change flow of African-Americans from white nonpoor to racially mixed neighborhoods is .194 (.194 + .082 = .275, shown in table 4). Therefore, about three-quarters of the neighborhood change flow of African-Americans out of white nonpoor neighborhoods is because neighborhoods become racially mixed. Many African-Americans who live in white nonpoor neighborhoods exit because their neighborhoods become blacker around them.

What do these patterns imply? The movement results indicate that *if* the neighborhoods nonpoor African-Americans are moving into would stay white and nonpoor after they move in, then the proportion of nonpoor blacks in white neighborhoods would rise considerably. But at the same time, some of the white nonpoor neighborhoods that had moderate concentrations of black families changed so that more than 30% of their popu-

lation was black, making them racially mixed neighborhoods. In other words, whites are moving out and blacks are moving in at a rate that is fast enough to keep the proportion of nonpoor blacks in white neighborhoods constant or slowly increasing, even though nonpoor African-Americans continue to have high net positive migration into white nonpoor areas.

The point estimates of neighborhood change for poor blacks in white nonpoor neighborhoods (−.072) also suggest neighborhood change moves poor blacks out of white nonpoor neighborhoods, but the flow is less substantial than for nonpoor blacks. This is largely because the number of poor blacks in white nonpoor neighborhoods is small to begin with.

As for nonpoor white neighborhoods, there is also a considerable net flow out of black nonpoor neighborhoods because of neighborhood change. Table 4 shows that .265% of the black nonpoor population moves out of black nonpoor neighborhoods each year due to neighborhood change. In a single year, a black nonpoor neighborhood can change to or from either a racially mixed neighborhood or a black poor neighborhood. We can, then, break down the net neighborhood change flow of African-Americans out of black nonpoor neighborhoods into flows into or out of racially mixed and black poor neighborhoods (shown in appendix table A1). There is a net neighborhood change flow of .172 into black nonpoor neighborhoods from racially mixed neighborhoods. Some racially mixed neighborhoods become black nonpoor neighborhoods. On the other hand, there is a population flow out of black nonpoor neighborhoods of .436 due to neighborhoods becoming poorer. Overall, the net neighborhood flow (difference) is .265 out of black nonpoor neighborhoods.

This suggests that nonpoor black neighborhoods often do not stay nonpoor but tend to lose nonpoor and gain poor residents until more than 20% of their population is poor. Many nonpoor black neighborhoods are in transition toward becoming black moderately poor neighborhoods. Table 4 shows correspondingly positive flows of nonpoor blacks into moderately (.253) and extremely poor (.176) black neighborhoods because of neighborhood change.

The patterns of neighborhood change for whites are also shown in the bottom portion of the two right-most columns of table 4. Neighborhood change appears to move nonpoor whites out of white nonpoor tracts (−.191) and into poor (.096) and racially mixed tracts (.087). This direction of change is similar to the pattern for African-Americans, but the size of the population flows, again, is smaller for whites.

For both blacks and whites, table 4 shows that neighborhood change around respondents appears to increase the share of the population in black and poor neighborhoods. The flows resulting from neighborhood change, however, are much larger for blacks than whites. There is no

evidence that neighborhood change around PSID respondents has moved substantial numbers of persons into nonpoor neighborhoods.

The neighborhood change results, combined with the results from the movement section, lead to two main conclusions: First, Wilson is correct that nonpoor African-Americans are moving in patterns that, if it was the only population movement, would increase their representation in white areas substantially—estimates suggest by about 30% in 10 years. But white flight and the increase in the black population in these neighborhoods are too rapid to allow for an overall increase in the proportion of the African-American population living in white nonpoor neighborhoods. Massey and Denton's finding that high SES African-Americans are only slightly more likely to live in white neighborhoods than low SES African-Americans, and Wilson's contention that nonpoor blacks are moving into white neighborhoods, then, are both supported by the data. Second, the proportion of the population residing in poor neighborhoods has increased substantially because of neighborhood change. This is especially the case for the African-American population. Combined with the earlier result that nonpoor persons migrate away from poor neighborhoods, this suggests that migration of the nonpoor away from the poor is a key mechanism leading to increases in the number of poor neighborhoods. Neighborhoods tend to deteriorate more often than they gentrify; the predominant path to a less poor neighborhood is to move into it.

## Are Poverty Rates among Stayers in Poor Neighborhoods Increasing?

The spatial mismatch and deindustrialization hypotheses imply that we should observe increases in poverty rates among residents of poor neighborhoods because the decline in inner-city manufacturing and the increasing frequency of spatial mismatches between jobs and people worsened the job prospects of inner-city workers during the 1970s and 1980s. Flows relevant to this hypothesis are presented in table 6. Unlike table 4, here the results are broken down into three periods, 1970–74, 1979–84, and 1985–90. I showed the pooled results in table 4 because the results for movement are generally the same across time periods; the flows into and out of poverty are not. Table 6 shows the net flow into (positive numbers) or out of (negative numbers) poverty among stayers in different neighborhood types. Again, the number can be interpreted as a portion of the total population by race, so the positive .020 indicates that on average .02% more of the black population lived in a white nonpoor neighborhood and entered poverty per year than exited it.

There is no systematic tendency in the data for more entrances into poverty than exits from poverty among residents of poor neighborhoods. In fact, for 1970–74 and 1985–90, poverty rates among African-American

TABLE 6

Average Net Population Movement into Poverty by Period and Neighborhood Type, Consecutive Stayer Person-Years Only

|  | 1970–74 | 1979–84 | 1985–90 |
|---|---|---|---|
| **Blacks:** | | | |
| White nonpoor | .020 | .046 | .069* |
| | (.1249) | (.0590) | (.0288) |
| Racially mixed | −.074 | .246 | −.159 |
| | (.1051) | (.1783) | (.1503) |
| Black nonpoor | −.069 | .018 | −.064 |
| | (.1439) | (.0415) | (.1206) |
| Black moderately poor | −.142 | .466* | −.131 |
| | (.2355) | (.2153) | (.1135) |
| Black extremely poor | −.168 | −.054 | −.241 |
| | (.1346) | ( 1524) | (.1475) |
| **Whites:** | | | |
| White nonpoor | −.192* | .191* | −.083* |
| | (.0952) | (.0389) | (.0349) |
| Racially mixed | −.030 | −.003 | −.021 |
| | (.0244) | (.0251) | (.0197) |
| White moderately and extremely poor | −.036 | .031 | −.029 |
| | (.0370) | (.0244) | ( 0308) |
| Racially mixed | −.030 | −.003 | −.021 |
| | ( 0244) | (.0251) | (.0197) |

Note —The figures are given in percentages of the total black population (top half) or white population (bottom half). So a .5 in the top half, e.g., represents .5% more of the black population entering than exiting poverty and living in the given neighborhood type. Positive numbers indicate net flows into poverty; negative numbers indicate net flows out of poverty. SEs are in parentheses.
* = flow/SE > 2

stayers in poor neighborhoods appear to be declining, although the decline is not statistically significant. Thus, there does not appear to be any consistent increase in the poverty rate among stayers in poor neighborhoods over time.[21] This is inconsistent with the idea that either growing spatial mismatch or increasing deindustrialization of industry had a strong, consistent impact on the poverty rates of residents of extremely poor neighborhoods for 1970–90.

This finding is not, however, conclusive evidence against the deindustrialization or spatial mismatch hypotheses. Nonpoor blacks might be moving into white nonpoor neighborhoods in part because they are following job prospects, in which case migration might be the proximate

[21] By a "stayer," I mean a consecutive person year in which a respondent does not change tract type.

cause of an increase in black poverty rates but changing economic opportunities would be the ultimate cause. Firmer tests of these theories will require data on the locations of jobs relative to people or interview data on why respondents are moving from one neighborhood to another.

Rather than conforming to a monotonic trend, the poverty-rate changes among stayers in table 6 is consistent with overall change in U.S. poverty rates. For most neighborhood types, the poverty change coefficients for 1970–74 and 1985–90 indicate decreases in the number of stayers who are poor (negative numbers), while those for 1979–84 indicate increases in the number of stayers who are poor (positive numbers). The individual coefficients are not usually statistically significant, but the result is quite consistent across the different neighborhood types. This corresponds basically to changes in the poverty rate in the United States during these periods: The poverty rate was declining or about holding steady in 1970–74 and 1985–90, while the poverty rate was increasing from 1979 until about 1983 (Lamison-White 1997).

Research has consistently found that black unemployment and poverty rates fluctuate more with business cycles than do white unemployment and poverty rates (Jaynes and Williams 1989; Freeman 1991), but it has not compared the experiences of persons living in different neighborhood types. Table 6 provides some evidence that residents of poor and black neighborhoods may have disproportionately born the brunt of the increase in poverty rates during the 1979–84 period. In table 6, there is a substantial increase in the number of poor African-Americans living in black, moderately poor neighborhoods during the 1979–84 period—on average .466% more of the black population entered poverty than exited it per year. African-Americans in the other neighborhood types do not register a similarly large increase. This suggests that the increase in poverty rates in the early 1980s was larger among blacks living in black and moderately poor neighborhoods than blacks living in other neighborhood types.

The finding that the increase in poverty during the early 1980s appears to have been especially concentrated among African-Americans living in moderately poor black neighborhoods suggests that the increase in poverty rates in the early 1980s contributed to the increase in the number of high-poverty neighborhoods. In fact, the increase in poverty rates was spatially distributed in such a way as to increase the number of extremely poor neighborhoods by a large amount, because the biggest increase in the poverty rate was concentrated in neighborhoods that were already moderately poor to begin with. The residents of moderately poor neighborhood may be more likely than residents of nonpoor neighborhoods to be at the bottom of the labor queue and thus to be most likely to be influenced by cyclical variations in labor demand.

This is a possible reason why Jargowsky (1997) finds such a strong relationship between changes in a metropolitan area's overall poverty rate and the number of extremely poor neighborhoods. If changes in poverty rates tend to be disproportionately concentrated in neighborhoods with moderately high rates of poverty to begin with, then cyclical variations in a metropolitan area's economy will tend to be strongly related to the number of extremely poor neighborhoods.

### The Relative Importance of Migration and Poverty Status Changes

A final result comes from comparing the flows that we observe in table 6 to the flows that we observe in table 4. An important conclusion from table 4 is that neighborhood change has led to an increase·in the number of whites and, especially, blacks living in poor neighborhoods over time. This is true in all three time periods (1970–74, 1979–84, 1985–90). Net flows resulting from changes in poverty status (table 6), on the other hand, indicate that changes in poverty status among persons in moderately poor neighborhoods are likely to have increased the number of extremely poor neighborhoods substantially during the early 1980s recession but not at other times. Except during the early 1980s recession, movement and neighborhood change appear to have played a larger role in increasing the number of poor neighborhoods over time than changes in poverty status.

### Completing the Decomposition

Thus far I have considered flows of persons who move, flows of persons whose neighborhood changed type around them, and flows of persons who change poverty status. This leaves out persons who experienced both a change in poverty status and a change in neighborhood type between PSID interviews. Flows of these persons are shown in appendix table A2. It is a somewhat unusual event for a respondent to change both their neighborhood type and poverty status in the same year. As a result, the flows of persons who experience both changes in poverty status and neighborhood type in a given year tend to be small and statistically insignificant.

One exception to the small size of the flows is that there is a moderate and statistically significant flow of African-Americans who both move out of poverty and into white nonpoor neighborhoods in the same year (.084). If this flow is added to the flow of nonpoor African-Americans who move into white nonpoor neighborhoods without changing their poverty status, then this further strengthens the case that nonpoor blacks are moving into white nonpoor neighborhoods more often than they are moving out.

CONCLUSIONS

Some expected and some surprising results have emerged from this analysis.[22] The basic conclusions follow:

1. The transition probabilities in the PSID data indicate that blacks, especially nonpoor blacks, move into white nonpoor neighborhoods more often than they move out. The difference is large enough to substantially increase the presence of nonpoor blacks in white neighborhoods over time.
2. When flows into white neighborhoods are considered simultaneously with flows resulting from neighborhood change, however, the share of blacks living in white nonpoor areas remains constant or increases slightly. These results suggest that if neighborhoods would stay predominately white as nonpoor blacks move in, then the proportion of nonpoor blacks living in white nonpoor neighborhoods would increase. But white populations tend to drop as blacks move in, and they do so at a fast enough rate to keep the proportion of black families in predominately white neighborhoods from increasing.
3. The population in black and poor tracts is increasing primarily because neighborhoods tend to become poorer and blacker around their residents, especially their black residents, not because of net positive movement of persons into poor neighborhoods.
4. There is no indication in the PSID data that stayers in black or poor neighborhoods experienced increases in their poverty rates in the 1970s and 1980s, except during the early 1980s recession. During the early 1980s recession, increases in the poverty rate among the nonpoor were spatially concentrated in black moderately poor neighborhoods. Since these neighborhoods were already moderately poor to begin with, this suggests that increasing poverty rates in the early 1980s had a strong effect in increasing the number of extremely poor neighborhoods.

Probably the most significant empirical contribution of this article is to provide an explanation of the contradictory results of tests of Wilson's

[22] A suggestion that has been made to me on several occasions is that I should calculate the flows in this article separately for different regions of the country. Unfortunately, the PSID sample size is not sufficient to allow for this analysis. Greene (1991) shows that there is an association between increases in neighborhood poverty rates and population loss even in Western cities such as Los Angeles, suggesting that the basic migration patterns described here hold for predominately black and white poor neighborhoods in all regions (but do not hold for predominately Hispanic tracts). The patterns for poverty change are more likely to vary by region, given that different regions have fared quite differently from one another in recent business cycles.

black middle-class out-migration thesis. I find that nonpoor African-Americans are moving into white areas fairly rapidly, as Wilson suggests. But the numbers of nonpoor African-Americans in white and nonpoor areas have not increased much over time, as Massey and Denton (1993) have shown, because of the decline in white population in these neighborhoods.[23] When considered as part of a dynamic system, the movement of blacks into white nonpoor neighborhoods and high continuing rates of racial segregation are not mutually exclusive.

Because poor blacks have not been moving into white neighborhoods as quickly as nonpoor blacks, the migration of nonpoor blacks out has been one factor contributing to the growing concentration of poverty in urban areas. This does not necessarily contradict the result of Massey et al. (1994) that middle-class black migration out of black neighborhoods is less important than racial segregation (or for that matter other factors) in explaining the existence of concentrated urban poverty, because factors that are relatively unimportant in explaining cross-sectional variation may be relatively important in explaining change over time.[24]

The analysis here particularly focuses on patterns for African-Americans and the white response to them. The data here does not include a sufficient number of other minority groups to allow for an analysis, but I think it unlikely that all of the same patterns would apply. Two factors make the situation of African-Americans unique relative to other racial and ethnic groups. First, negative stereotypes about other racial and ethnic groups are often not as strong. Evidence suggests that whites are not likely to go to nearly as great a lengths to avoid coresidence with members of other racial and ethnic groups as they are to African-Americans (e.g., Massey and Denton 1987). Second, there is less migration into black urban neighborhoods from international destinations than is the case for Hispanic or Asian neighborhoods.

DISCUSSION

The results in this article suggest that migration of the nonpoor away from the poor has played a key role in forming new poor urban neighborhoods

---

[23] A likely cause of this decline in white neighborhood population is white flight, but it could also be accomplished by reductions in the number of whites moving in even if the number of whites moving out remains stable. It is also possible that whites are moving out largely because of characteristics that are correlated with large black neighborhood populations rather than a purely racial motivation (e.g., Harris 1997).

[24] Because theoretical models imply nonlinear and interactive relationships among these variables, it is difficult to measure their relative importance. Past research assessing the causes of high-poverty neighborhoods has often not fully grappled with this problem.

during the 1970s and 1980s. Increases in poverty rates among persons living in moderately poor neighborhoods also contributed significantly to the increase in the number of extremely poor neighborhoods but did so *only* during the early 1980s recession. Migration by the nonpoor away from the poor, on the other hand, has been an important factor in increasing the number of neighborhoods with high rates of poverty *throughout* the 1970s and 1980s.

The patterns of migration discussed here do not displace or contradict existing theories about poor neighborhoods, such as those developed by Wilson (1987, 1996) or Massey and Denton (1993). Instead, the results further specify their models by suggesting that migration has been a key mechanism through which the middle class and affluent have sought to segregate themselves from the poor, whites have sought to avoid black neighbors, and workers have adjusted to changes in urban job markets. The desire of the nonpoor to avoid poor neighborhoods and of whites to avoid black neighbors can be thought of as push migration factors motivating moves by the middle class and whites. The migration patterns described in this article are likely to have resulted from these push factors combined with three historical changes: the increasing ability of many African-American families to move into white neighborhoods, improvements in intraurban transportation, and the growing suburbanization of employment.

Before about 1970, racial violence against blacks in white neighborhoods and explicitly discriminatory practices in housing markets were extremely strong barriers faced by black families who might consider moving into a white neighborhood (Massey and Denton 1993). Racial disparities in the probability of moving into a white neighborhood that Massey et al. (1994) document and the results of the fair housing audits (Yinger 1995) demonstrate that these barriers remain significant. But the results here suggest that during the 1970s and 1980s these barriers were not so strong as to prevent the movement of significant numbers of blacks into white neighborhoods.[25] If it were not for the continuing movement of whites away from neighborhoods with increasing black concentrations, the proportion of the nonpoor black population living in white neighborhoods would have increased substantially.

[25] Studies of changes in neighborhoods before 1970 suggest the movement of blacks into white neighborhoods in the 1950s and 1960s largely reflects the enormous demand for housing in black areas relative to the limited supply caused by massive migration of blacks into urban areas (see Aldrich [1975] for a review) This led to much higher rents for housing in black areas than for comparable housing in white areas and made racial turnover potentially a very profitable proposition for landlords and real estate agents.

As barriers to blacks entering white neighborhoods have weakened, middle-class blacks have increasingly sought to avoid neighborhoods with high poverty rates by moving to middle-class and often white neighborhoods; whites have likewise sought to achieve segregation from blacks by moving out of neighborhoods with increasing black populations. Inner-city manufacturing jobs were once a strong factor pulling migrants into central cities, but the suburbanization of employment and improvements in transportation have meant this is no longer the case. The result has been the emptying out of poor neighborhoods. Neighborhoods in transition to high-poverty status empty first of whites, then of many middle-class blacks, leaving more disadvantaged and less populous areas. This is consistent with decennial census results that show that poor neighborhoods have been increasing in geographic size and falling in population density (Jargowsky and Bane 1991; Jargowsky 1997).

Compared with descriptions of poor neighborhoods in earlier historical periods, the results suggest an important historical change in how poor neighborhoods have formed. Accounts of racial succession of neighborhoods in the 1950s found that neighborhoods undergoing racial transition tend to increase in population density, especially in passing through a late phase in racial succession referred to as "piling up" in which previously white-owned homes and apartments are subdivided into smaller dwellings to accommodate the housing demands of black immigrants (Duncan and Duncan 1957). During the 1950s and 1960s, the pull of good inner-city jobs for low-skilled workers dominated, leading to the migration of large numbers of black and poor immigrants into inner-city neighborhoods. Although the affluent have always made efforts to segregate themselves from the poor, migration into cities was proceeding at too rapid a pace to allow inner-city neighborhoods to drop substantially in population as part of this process. With the decline of inner-city industrial employment and the corresponding end of black migration into urban areas, poor African-American neighborhoods have changed from densely packed communities of recently arrived immigrants to areas gradually abandoned by the nonpoor. The cessation of the flow of black immigrants to the nation's cities, and the corresponding decline in the population density of poor neighborhoods, may be one unexplored factor responsible for the change in the nature of poor African-American neighborhoods in the early 1970s that Wilson (1987) describes.

APPENDIX

TABLE A1

Flow Breakdowns for White and Black Nonpoor Neighborhoods by Origin
and Destination, African-American Respondents

| | Flow into Nonpoor Neighborhood (from Specified Type) | Flow out of Nonpoor Neighborhood (into Specified Type) | Net Flow into Nonpoor Neighborhood (Difference of Columns) |
|---|---|---|---|
| White nonpoor neighborhoods: | | | |
| Neighborhood change flow between nonpoor white neighborhoods and white poor neighborhoods ... . ... ... . | .002 (.0011) | .084 (.0358) | −.082 |
| Neighborhood change flow between nonpoor white neighborhoods and racially mixed neighborhoods ... ... | .056 (.0211) | .250 ( 1037) | −.194 |
| Average net flow (sum of differences, matches net flow shown in table 4) ... .. ......... | | | −.275 (.0748) |
| Black nonpoor neighborhoods: | | | |
| Neighborhood change flow between nonpoor black neighborhoods and racially mixed ... ... ... . .. ... .. ............. | .234 (.0485) | .062 (.0400) | 172 |
| Neighborhood change flow between nonpoor black neighborhoods and poor black neighborhoods ..... ..... ....... | 081 (.0446) | .517 (.1163) | −.436 |
| Average net flow (sum of differences, matches net flow shown in table 4) ... ......... ... | | | −.265 (.1300) |

Note —The flows are given as a percentage of the total black population. SEs are in parentheses
Differences do not sum to the total net flow due to rounding error

TABLE A2

AVERAGE NET POPULATION FLOW AMONG NEIGHBORHOOD TYPES,
CONSECUTIVE PERSON YEARS IN WHICH THE RESPONDENT CHANGED
POVERTY STATUS AND NEIGHBORHOOD TYPE, POOLED DATA FROM
1970–74 AND 1979–90

| Neighborhood Type | Nonpoor | Poor |
|---|---|---|
| Blacks. | | |
| White nonpoor ..................... . .. .. ..... ....... | .084* | 003 |
| | (.0127) | (.0257) |
| Racially mixed ............................ .. | −.026 | −.037 |
| | ( 0399) | (.0211) |
| Black nonpoor .............................. ... .. .. . .. .. | −.007 | −.037 |
| | ( 0368) | (.0321) |
| Black moderately poor ................................ | .028 | .025 ` |
| | (.0434) | (.0182) |
| Black extremely poor . _ .. ... .. . ... .. ...... | −.026 | .039 |
| | (.0093) | (.0244) |
| Whites· | | |
| White nonpoor .......... ..... ... .. .. ... .. .. .. . | −.033* | .010 |
| | (.0064) | (.0150) |
| White moderately and extremely poor ... .. | .013 | .014* |
| | (.0124) | (.0060) |
| Racially mixed ...................... .. .. .. .. .. ..... | .001 | −.000 |
| | ( 0039) | (.0038) |

NOTE —The flows are given as a percentage of the total white or black population, respectively. See notes to table 4 for an explanation of population figures SEs are in parentheses
* = flow/SE > 2

REFERENCES

Aldrich, Howard. 1975. "Ecological Succession in Racially Changing Neighborhoods: A Review of the Literature." *Urban Affairs Quarterly* 10 (3): 327–48.

Boudon, Raymond. 1993. *Mathematical Structures of Social Mobility*. San Francisco: Jossey-Bass.

Denton, Nancy A., and Douglas S. Massey. 1988. "Residential Segregation of Blacks, Hispanics, and Asians by Socioeconomic Status and Generation." *Social Science Quarterly* 69 (4): 797–817.

Duncan, Otis Dudley, and Beverly Duncan. 1957. *The Negro Population of Chicago: A Study of Residential Succession*. Chicago: University of Chicago Press.

Farley, Reynolds. 1991. "Residential Segregation of Social and Economic Groups among Blacks, 1970–80 " Pp. 274–98 in *The Urban Underclass*, edited by Christopher Jencks and Paul E. Peterson. Washington, D.C.: Brookings Institution.

Farley, Reynolds, and William Frey. 1994. "Changes in the Segregation of Whites from Blacks during the 1980s: Small Steps toward a More Integrated Society." *American Sociological Review* 59:23–45.

Freeman, Richard. 1991. "Employment and Earnings of Disadvantaged Young Men in a Labor Shortage Economy." Pp. 103–21 in *The Urban Underclass*, edited by C. Jencks and P. Peterson. Washington, D.C.: Brookings Institution.

Gramlich, Edward, Deborah Laren, and Naomi Sealand. 1992. "Moving into and out of Poor Urban Areas " *Journal of Policy Analysis and Management* 11 (2): 273–87.

Greene, Richard 1991. "Poverty Area Diffusion: The Depopulation Hypothesis Examined." *Urban Geography* 12 (6) 526–41.

Harris, David R. 1997. "'Property Values Drop when Blacks Move In, Because . .': Racial and Socioeconomic Determinants of Neighborhood Desirability." Research Report no 97-387. Population Studies Center, Ann Arbor, Mich.

Harrison, Bennett, and Barry Bluestone. 1981. *The Great U-Turn: Corporate Restructuring and the Polarizing of America.* New York: Basic Books.

Hill, Martha. 1992. *The Panel Study of Income Dynamics: A User's Guide.* Newbury Park, Calif.: Sage.

Holzer, Harry J. 1994. "Black Employment Problems: New Evidence, Old Questions." *Journal of Policy Analysis and Management* 13 (4): 699–722.

Hughes, Mark Alan 1990. "Formation of the Impacted Ghetto: Evidence from Large Metropolitan Areas: 1970–1980." *Urban Geography* 11:265–84.

Jakubs, John F. 1986 "Recent Racial Segregation in the U.S. SMSAs." *Urban Geography* 7:146–63.

Jargowsky, Paul A. 1994 "Ghetto Poverty among Blacks in the 1980s." *Journal of Policy Analysis and Management* 13 (2): 288–319.

——— 1996. "Take the Money and Run: Economic Segregation in U.S. Metropolitan Areas." *American Sociological Review* 61:984–98

———. 1997. *Poverty and Place: Ghettos, Barrios, and the American City.* New York: Russell Sage Foundation.

Jargowsky, Paul A., and Mary Jo Bane. 1991. "Ghetto Poverty in the United States, 1970–1980." Pp. 235–73 in *The Urban Underclass,* edited by Christopher Jencks and Paul E. Peterson. Washington, D.C.: Brookings Institution.

Jaynes, Gerald David, and Robin M. Williams, eds. 1989. *A Common Destiny: Blacks and American Society.* Washington, D.C.: National Academy Press.

Jencks, Christopher, and Susan Mayer. 1990. "Residential Segregation, Job Proximity, and Black Job Opportunities." Pp. 187–222 in *Inner-City Poverty in the United States,* edited by Lawrence E. Lynn and Michael McGeary. Washington, D.C. National Academy Press.

Kain, John F. 1968. "Housing Segregation, Negro Employment, and Metropolitan Decentralization." *Quarterly Journal of Economics* 82:175–97.

Kalton, Graham. 1979. "Practical Methods for Estimating Survey Sampling Errors." *Bulletin of the International Statistical Institute* 47 (3): 495–514.

Kasarda, John. 1990. "Urban Industrial Transition and the Underclass." *Annals of the American Academy of Political and Social Science* 501:26–47.

Keyfitz, Nathan. 1977. *Applied Mathematical Demography.* New York: John Wiley & Sons.

Korenman, Sanders, John Sjaastad, and Paul A. Jargowsky. 1995. "Explaining the Rise of Ghetto Poverty in the 1980s: A Replication and Extension." Manuscript. University of Minnesota, Humphrey Institute for Public Affairs.

Krivo, Lauren, and Ruth D. Peterson. 1996. "Extremely Disadvantaged Urban Neighborhoods and Urban Crime." *Social Forces* 75 (2): 619–50.

Lamison-White, Leatha. 1997. "Poverty in the United States: 1996." Current Population Reports, Series P60-198. Washington, D.C.: Government Printing Office.

Massey, Douglas S. 1979. "Effects of Socioeconomic Factors on the Residential Segregation of Blacks and Spanish Americans in U.S. Urbanized Areas." *American Sociological Review* 44:1015–22.

———. 1990. "American Apartheid: Segregation and the Making of the Underclass." *American Journal of Sociology* 96.329–57.

Massey, Douglas S., and Nancy A Denton. 1987. "Trends in the Residential Segrega-

tion of Blacks, Hispanics and Asians." *American Sociological Review* 52·802–25.

————. 1993. *American Apartheid: Segregation and the Making of the Underclass.* Cambridge, Mass.: Harvard University Press

Massey, Douglas S., and Mitchell L. Eggers. 1990. "The Ecology of Inequality: Minorities and the Concentration of Poverty, 1970–1980." *American Journal of Sociology* 95:1153–89.

————. 1993. "The Spatial Concentration of Affluence and Poverty during the 1970s." *Urban Affairs Quarterly* 29 (2): 299–315.

Massey, Douglas S., Andrew Gross, and Kumiko Shibuya. 1994. "Migration, Segregation, and the Geographic Concentration of Poverty." *American Sociological Review* 59:425–45.

Morenoff, Jeffrey D., and Robert J. Sampson. 1997. "Violent Crime and the Spatial Dynamics of Neighborhood Transition: Chicago, 1970–1990." *Social Forces* 76 (1): 31–64.

Preston, Samuel H., and Cameron Campbell. 1993. "Differential Fertility and the Distribution of Traits: The Case of IQ." *American Journal of Sociology* 98 (5): 997–1019.

Rodgers, Joan R., and John L Rodgers. 1993. "Chronic Poverty in the United States." *Journal of Human Resources* 28 (1): 25–53.

Rogers, Andrei. 1968. *Matrix Analysis of Interregional Population Growth and Distribution.* Berkeley: University of California Press.

Skogan, Wesley. 1990. *Disorder and Decline: Crime and the Spiral of Decay in American Neighborhoods.* Berkeley and Los Angeles: University of California Press.

————. 1997. *Stata Statistical Software Release 5.0 User's Guide.* College Station, Tex.: Stata Press.

South, Scott J., and Kyle D. Crowder. 1997. "Escaping Distressed Neighborhoods: Individual, Community, and Metropolitan Influences." *American Journal of Sociology* 102 (4): 1040–84.

Survey Research Center, Economic Behavior Program. 1998. *A Panel Study of Income Dynamics: 1968–1993 Individual Data Codebook.* University of Michigan, Inter-University Consortium for Political and Social Research.

Teitz, Michael, and Karen Chapple. 1998. "The Causes of Inner-City Poverty: Eight Hypotheses in Search of Reality." *Cityscape* 3 (3): 33–70.

U.S. Bureau of the Census. 1983a. Census of Population and Housing, 1980: 1980–1970 Tract Comparability File. Washington, D.C.: Bureau of the Census.

————. 1983b. Census of Population and Housing, 1980: 1980–1970 Tract Comparability File Technical Documentation. Washington, D.C.: Bureau of the Census.

————. 1992a. Census of Population and Housing, 1990: Tiger/Census Tract Comparability File. Washington, D C.: Bureau of the Census.

————. 1992b. Census of Population and Housing, 1990: Tiger/Census Tract Comparability File Technical Documentation. Washington, D C.: Bureau of the Census.

————. 1995. Poverty Areas. Statistical Brief 95-13. Washington, D.C.: Bureau of the Census.

White, Michael J. 1987. *American Neighborhoods and Residential Differentiation.* New York: Russell Sage

Wilson, William Julius. 1987. *The Truly Disadvantaged.* Chicago. University of Chicago Press.

————. 1996 *When Work Disappears: The World of the New Urban Poor.* New York: Alfred A. Knopf.

Wolter, Kirk M. 1985. *Introduction to Variance Estimation* New York: Springer-Verlag.

Yinger, John. 1985. *Closed Doors, Opportunities Lost. The Continuing Costs of Housing Discrimination.* New York: Russell Sage.

# How Events Enter the Public Sphere: Conflict, Location, and Sponsorship in Local Newspaper Coverage of Public Events[1]

Pamela E. Oliver
*University of Wisconsin, Madison*

Daniel J. Myers
*University of Notre Dame*

Protest events occur against the backdrop of public life. Of 382 public events in police records for one year in a small U.S. city, 45% convey a message, 14% involve social conflict, and 13% are standard protest event forms. Local newspapers covered 32% of all events, favoring events that were large, involved conflict, were sponsored by business groups, and occurred in central locations. The more liberal paper also favored rallies and events sponsored by national social movement organizations (SMOs) or recreational groups. Discussion centers on the ways these factors shape the content of the public sphere.

## INTRODUCTION

Public life happens in public places. In public, people who are not intimates may meet face to face and perhaps influence one another or a wider public. Street theater, street-corner speeches, marches, celebrations, vigils, leafleting, and other kinds of public acts seek to express collective sentiments or influence public opinion. Scholars around the world recognize the importance of public events for the public sphere, the abstract space in which citizens discuss and debate public issues (e.g., Alario 1994; Chaffee 1993; Koenen 1996; Sebastiani 1997). The link between public events and the public sphere is the mass media. Acts staged in public places may seek to influence only the other people in that space: passersby who will hear the speech, song, or chant, or read the signs. But, more often, these public events are oriented not only to those physically present

but also to a larger public. Usually, a major goal of a public event is to attract the attention of the mass media, for only through the mass media can people communicate beyond their immediate social setting.

Scholars of protest have long recognized that the functioning of the mass media has critical consequences for social movements and cycles of protest, generally agreeing with Lipsky: "If protest tactics are not considered significant by the media, or if newspapers and television reporters or editors decide to overlook protest tactics, protest organizations will not succeed" (1968, p. 1151). The mass media play a critical role in the progression of a protest cycle, communicating information about events and issues in "media attention cycles" (Downs 1972) that first attract attention to developing events and issues and then withdraw attention from them when the issue becomes old or the disruption too severe (Tarrow 1994, 1998). Many scholars have stressed the importance of the media for shaping the public perception and framing of movements and movement issues (Gamson and Modigliani 1989; Gamson and Wolfsfeld 1993; Gans 1980; Hilgartner and Bosk 1988; Kielbowicz and Scherer 1986; Molotch 1979; Molotch and Lester 1974; Olien, Tichenor, and Donahue 1992; Oliver 1989; Zald 1996). Understanding the filter applied by newspapers and television to the realities of demonstrations is essential to understanding the effects of protests and demonstrations in the polity.

Apart from their crucial role in shaping protest cycles, the media are at least as pivotal in research on protest and unrest because scholars so often turn to media sources (particularly newspapers) for data about protest events. (A few influential examples are Gurr [1968], Jenkins and Eckert [1986], Jenkins and Perrow [1977], Kriesi et al. [1995], Lieberson and Silverman [1965], McAdam [1982], Olzak [1992], Shorter and Tilly [1974], Spilerman [1970, 1976].) The validity of these studies depends heavily on the accuracy and completeness of newspaper records or at least on the assumption that the sample of events covered by newspapers reliably tracks the underlying flow of events according to a consistent principle such as random selection, possibly with probability proportionate to the "size" of the event.

Unfortunately, this methodological assumption treats the media as passive "channels" of communication or neutral and objective observers and recorders of events, a view that for some time now been rejected by scholars of the media (e.g., Gans 1980; Herman and Chomsky 1988; Shoemaker and Resse 1991), as well as refuted by studies of the media coverage of collective events (Danzger 1975; Franzosi 1987; McCarthy, McPhail, and Smith 1996; Mueller 1997a; Snyder and Kelly 1977). Pinning down the nature of the bias is difficult, however, because independent tabulations of protest events against which to compare the media record are extraordinarily rare. In recent years, however, a new tactic has reinvigorated the

study of bias in media records of protest. Police department records compiled routinely provide an alternate source that can be compared to media records. Police records contain many more events than are recorded in the press, which allows patterns of selection bias in the media record to be better identified (Fillieule 1998; Hocke 1998; McCarthy, McPhail, and Smith 1996; McCarthy et al. 1998).

However, these studies of police and media records of protests have, themselves, been distorted by a failure to examine media representations of protest in the context of comparisons with other kinds of public actions. Without such comparisons, scholars have been unable to provide a complete account of the role of protest or movement events within the public life of a community. This article addresses this lapse by presenting a comprehensive picture of the public events in one small city and showing how protest events fit within this larger context. We show that, although not all protest events are covered by local newspapers, conflictual events are generally *more* likely to receive media coverage than other events of similar size and form, and that space itself (i.e., a central location) plays a much greater role in media coverage than has previously been recognized. These results suggest a need to reconsider the interplay between public events and the mass media in the creation of public discussions about social issues.

Public Acts and Public Messages in Public Spaces

Although it is beyond the scope of this article to engage theoretical debates about the nature of the public sphere, we take it as given that the public discussion of public issues is important, that public events play a significant role in promoting or shaping the public sphere, and that mass media are important conduits of information between groups of people. Understanding the way public events become represented in the mass media is critical for understanding the workings of the public sphere.

It is well recognized that public life is ritualized and organized, so that public actions generally take on well-defined forms that are meaningful to participants and observers (Bennett 1980; Brown and Kimball 1995; Drucker and Gumpert 1996; Lees 1994). Insiders may accept these forms as natural categories of action, while detached observers can observe the ways in which the categories themselves are constructed and evolve over time. The "protest" is one such ritualized form that conveys roughly the same meaning to activists, police, news reporters, the general public, and social scientists alike. This shared meaning has blinded researchers to the constructed nature of "protest" and led them to assume an unproblematic isomorphism between form and content in their definitions of protest events. But, as Tilly (1978) first told us 20 years ago, the forms or reper-

toires of protest shift across time and space, and new forms of protest are often created by adapting nonprotest forms to new purposes. Identical forms may carry very different content. "Parade" and "march" are two names for exactly the same form (McPhail and Wohlstein 1986), and the words can be used interchangeably even though in the United States in the 1990s the popular connotations of *parade* involve entertainment, while the word *march* is popularly applied only to message events. Likewise, there are many kinds of rallies, from pep rallies to protest rallies: they share the form of a stationary gathering with speeches containing informational and emotional content, but vary greatly in the "issue" they may address. As protest repertoires evolve, message content is often added to event types created for other purposes. In the United States in the 1990s, ceremonies, musical performances, literature distribution, and amateur street theater are all event types that are typically "apolitical," but all have carried protest content in past times and places and can and do sometimes carry protest content in the 1990s.

At any given time, there are certainly regularities and patterns about the kinds of content conveyed in various forms, but these meanings are always contextual and always evolving. Block parties are generally understood in the United States in this era to be consensual events that convey a sense of sociability and community to residents of a particular area, although they do disrupt normal traffic patterns because they require closing a street. However, using a barricade to close a street without permission has been internationally recognized as a protest tactic since the French Revolution. In the context of ongoing student-police battles in the late 1960s, when countercultural residents blocked Madison's Mifflin Street for a block party, the police treated the event as an insurrection, and the ensuing battle between police and residents became a full-scale riot. Even in the apolitical 1990s, the Mifflin Street block party evolved into a riot in 1995, although the drunken revelers setting fires and attacking police had no apparent political agenda.

Similarly, we can assess biases in the coverage of protest or claims-making content only by measuring it against the coverage of other kinds of content or the lack of message content. Social movements and protest events researchers have generally focused their attention on particular kinds of claims or particular kinds of events, on claims or actions that are contentious, disruptive, political, or directed at the government (e.g., Fillieule 1998; McCarthy, McPhail, and Smith 1996; Olzak, Shanahan, and West 1994; Rucht and Neidhardt 1998; Tilly 1986). Such restrictions, however, ignore the possibility that the same issue can be constructed as contentious or not and that the boundary between contentious and consensual issues is always evolving. Both AIDS and breast cancer, for example, have been constructed both as health concerns that should be dealt

with consensually and as contentious issues about public funding for medical research. The Freedom from Religion Foundation has challenged the consensual nature of Christmas trees and nativity scenes on public property. Disruption always leads to public debate about whether it is meaningful protest or meaningless hooliganism. To study only issues framed as political or conflictual is to be unable to examine the effect of the frame itself on the reception of the message.

To obtain a wide range of events for comparison, we used police records to compile a broad set of public collective events, both contentious and consensual. As we indicate below, even this procedure could not catch "all" events, because police agencies themselves are selective in which events they record, but we have obtained a broad set of events compiled from sources other than the media that may be used for identifying the selection factors in media coverage. We then compare the media coverage for different types of events. We classify an event as having a "message" if part of its apparent purpose was to influence public opinion or action, broadly construed. A message in this sense encompasses protest messages or claims but also includes consensual persuasive content (e.g., breast cancer awareness), ethnic pride, religious messages, fund-raising, and business promotion. Events that do not have messages are those that are ends in themselves, such as social events, performances, and athletic competitions. There are obviously ambiguous cases in informational or educational events, particularly when incomplete records are being coded. The criterion for identifying a message, however, was whether the event appeared to be an end in itself that did not need any recognition from nonparticipants to accomplish its purpose (no message) or whether the event was a means to exert influence over persons not participating (message). We defined an event as involving *conflict* if we were aware of social or political conflict about the event itself, the sponsoring organization, or the issue addressed by the event. One event, the Mifflin Street block party, was judged to involve conflict despite having no apparent message because of the history of the event and substantial public debate about it. Otherwise, conflict events were a subset of message events.

Local News

Although most research attention over the years has been focused on national newspapers of record, scholars have begun to recognize the importance of local and regional news media, both substantively, for their role in shaping local political discourses, and methodologically, for their relative comprehensiveness in covering local events. More people read local newspapers than read the *New York Times*. Many events that come to be perceived as national news, or as part of a national movement, begin as local

events that make the local news. A full understanding of the role of the media in shaping protest cycles requires an examination of the construction of local news and the ways local politics and local protest are intertwined with national movements and national politics. Studies of local protest campaigns, particularly those that have compared similar campaigns in different locales, have found differences that are tied to the particular social, political, and economic conditions of an area; studies of local protest reveal the influence of national or global movements as they play out in a particular way in each particular place (e.g., Eisinger 1973; Hellman 1987; Ray 1993; Rothman 1993). Although the most local newspapers are small suburban advertising weeklies, newspapers classed as "regional" sources provide serious news coverage of their metropolitan areas. Regional news media cover a much higher proportion of the events within their catchments than do national media, and as electronic archives of newspapers make searches of a compilation of regional newspapers feasible, scholars are recognizing that a collection of regional newspapers may provide a much more comprehensive documentation of events than any national newspaper ever could. However, more studies are needed of the construction of local news before this potential can be realized.

## THE ROUTINES OF PROTEST AND PUBLIC EXPRESSION

As Mueller argues in a very useful review (1997b), it is important to work toward a theory of media production that can lead to an understanding of which events do and do not become news. Central to such a theory is a realistic understanding of the production of public events. Most research and theory about the media coverage of events are implicitly based on a neat image of the relation between protesters, the police, and the media: activists plan and carry out an event, police are caught by surprise and respond to the event, and the media hurry to the event so they can observe it and report what happened in the confrontation. But recent scholarship has shown that this neat image is more wrong than right, at least most of the time for the 1990s in the United States and much of Europe. Experienced event organizers, police, and reporters are more like members of an improvisational troupe: the script is not fixed, but the players have worked together before, they follow general guidelines, and they can predict each other's actions. Most public events—often even the majority of protests—have permits, and the organizers and police jointly agree upon the time, place, and manner of the event. Even in unpermitted disruptive protests, experienced protesters and police often negotiate the terms of the event, agreeing on what conduct will or will not result in an arrest. The media are similarly tied up in the creation of the event. Experienced organizers plan events that meet journalistic standards of newsworthiness,

43

write press releases, call reporters, and craft their "sound bites" for the media; in some cases, the police complain that reporters are notified about unpermitted protests before the police are. Reporters cultivate activists as "sources," and activists cultivate reporters in the hopes of gaining more favorable coverage for their issue. In short, understanding the routine interplay of activists with police and reporters is an essential precondition for understanding "media bias."

Policing Public Events

Free speech decisions in U.S. courts during the 1970s determined that authorities could not regulate the content of public expression but that they could regulate its time, place, and manner. As a result, police agencies began instituting standardized permitting procedures (McCarthy and McPhail 1997; McCarthy, McPhail, and Crist 1998; McPhail, Schweingruber, and McCarthy 1998). Although operating in different national legal climates, the trend in Europe has been in the same general direction, away from confrontation and toward regulation and negotiation (Della Porta 1996a; Della Porta 1996b; Della Porta and Reiter 1998). From some points of view, the result has been a win-win situation. The police gain knowledge about what will happen and are better able to minimize disruption related to the event. Activists can count on police help to divert traffic, manage crowds, and ensure no one gets hurt or arrested. Although there are periods of intense mobilization during which these arrangements break down and times in which being arrested becomes a badge of honor for an activist cadre, most people most of the time do not want to spend time in jail or to pay large fines. Similarly, police agencies do not wish to police riotous events or to appear to have lost control of a protest situation. The result is that the majority of political protest events stay within the bounds of time, place, and manner restrictions.

As a consequence, police and protest practices have evolved together. These mutual adjustments between the authorities and the protesters have shaped the recent cycles of protest in the United States and Western Europe (e.g., Costain 1992; Koopmans 1993; McAdam 1983; McCarthy, Britt, and Wolfson 1991; McCarthy and McPhail 1997; McPhail, Schweingruber, and McCarthy 1998; Meyer 1993; Tarrow 1994). One aspect of this developing interaction is the evolution of successful protest campaigns toward increasingly formalized protest organizations that in turn shift their actions toward more institutionalized actions directed by professional staff (McCarthy and McPhail 1997; McCarthy and Zald 1977; Oliver and Marwell 1992; Staggenborg 1988; Tarrow 1994). Another consequence is that protest forms diffuse from hard core activists to the broader population, so that a rally or march is a relatively legitimate form of action

in the eyes of the general public and, thus, can carry many different kinds of content. Although the traditional conception of social movements emphasizes protest as an extrainstitutional form of political behavior, it is clear that protests and demonstrations have themselves become ritualized and institutionalized (e.g., Lofland and Fink 1982; Oliver and Marwell 1992).

Media Filters and Event Coverage

Although there is a casual tendency to treat the lack of coverage of any event as "media bias," no one seriously argues that it is possible for newspapers to report every event that happens, and even those who defend the concept of "objective" news reporting say that there are standards of "newsworthiness" that determine which events are "news" and which are not. The task therefore is to identify the systematic factors that determine the likelihood that an event will receive news coverage. Three kinds of factors have been identified as influencing the news coverage of events: the predispositions of news organizations or of particular reporters toward certain kinds of events or issues, journalistic norms and standards for assessing the news value of events and issues, and the mundane routines of producing news reports to deadlines.

*Predispositions.*—Movement-oriented commentators often stress the predispositions of news organizations or reporters in selecting events for coverage. These emphasize the increasingly concentrated control of the mass media (Bagdikian 1983; Downing 1980; Lee and Solomon 1990), cite instances in which coverage of events has been downplayed or omitted entirely because of concerns about social disorder (e.g., Gitlin 1980; Herman and Chomsky 1988; Molotch 1979; Parenti 1986), and argue that news media generally downplay union or labor struggles (e.g. Beharrell and Philo 1977; Hartmann 1979; Morley 1976). It is also well established that more powerful people and institutions have more ready access to the media (e.g., Goren 1980; Shoemaker and Resse 1991). Nevertheless, comparisons among specific news organizations often find that the overt editorial policies of a newspaper find expression in the selection of events that receive attention in the news sections and particularly that more left-wing newspapers cover more movement-related events (e.g., Franzosi 1987; Kriesi et al. 1995, p. 256).

*News value.*—Journalistic norms and standards for assessing the news value of events are widely agreed to be important in determining which events get covered (Altheide and Rasmussen 1976; Bowers 1967; Gamson and Meyer 1993; Gamson 1995; Gamson et al. 1992; Gans 1980; Hocke 1998; Kielbowicz and Scherer 1986; Schulz 1982; Shoemaker and Resse 1991; Sigelman 1973; Snyder and Kelly 1977). A standard prescriptive list

of news value criteria taught to journalism students generally includes the prominence or importance of the issue including the number of people affected and the magnitude of the effect; human interest and human drama; conflict or controversy; the unusual; timeliness; and proximity, with a preference for local events over distant ones (Shoemaker and Resse 1991). Working journalists and media researchers studying them agree that these news value criteria are complex and competing and that a story's news value is constructed in the process of making assignment decisions and writing stories (Lester 1980; Molotch and Lester 1974). Reporters see or create an angle for their "story" and in the process construct subjective news value features such as "narrativity" (Jacobs 1996) and "drama" (Gamson and Meyer 1996). These news values are understood by many activists who seek to create events that will embody them (Cohn and Gallagher 1984; Gamson and Wolfsfeld 1993; Ryan 1991; Salzman 1998).

As Hocke (1998) emphasizes, however, empirical attempts to tie news value to events identify different characteristics, and assessments of the criteria of news value typically are made on the basis of the features of published reports, not on the inherent character of the original events prior to their construction as "news." The only consistent claim validated through comparing media coverage to extramedia sources is that events that are "bigger" in terms of involving more people, lasting longer, or creating more disruption are more likely to be covered than "smaller" events (Kriesi et al. 1995; McCarthy, McPhail, and Smith 1996; Snyder and Kelly 1977).

*News routines.*—Although news value is often folded into the concept of "news routines," it is helpful to distinguish the more mundane news routine constraints on the reporter's job, specifically the problems of getting information and writing to meet a deadline. The constraints of a deadline create a premium for stories that fit into an obvious template. Reporters cultivate "sources" who will give them information for their stories, and groups and individuals differ markedly in their level of routine access to reporters (Shoemaker and Resse 1991). Other mundane factors are well recognized by media analysts and activists, and include writing a good press release with vivid quotations that can be incorporated into a news story, timing the event appropriately for news media deadlines, cultivating relationships with reporters, notifying the press in advance of upcoming events, and (for television coverage) planning events for their visual appeal (Cohn and Gallagher 1984; Gamson and Wolfsfeld 1993; Ryan 1991; Salzman 1998).

Considerations of routine suggest the importance of other factors that have been less well recognized in the literature on protest events. Although spontaneous protests are in some sense more newsworthy because they

46

are more likely to be disruptive, they are much harder for reporters to cover. Events that are announced in advance permit reporters to plan their schedules. Annual events are even more amenable to planning and permit the preparation of feature or human-interest stories that can be pegged to the event. The location of an event matters, too. Past research has emphasized "national" media and has demonstrated that geographic proximity of the event to the media outlet is an important factor in coverage (Danzger 1975; Hocke 1998; Mueller 1997*a;* Snyder and Kelly 1977), which is quite consistent with the news value of proximity. However, the importance of proximity as a news routine factor has been less well recognized for coverage within a city, with little attention having been paid to the spatial dimension of news "beats": the "state government" beat occurs at the state capitol, the "local government" beat at city hall, the "crime" beat at the police station. Events occurring in places frequented by reporters in their daily routines are much easier for them to cover than those occurring elsewhere.

## METHODS AND DATA SOURCES

### Setting

Madison, Wisconsin, is a city of about 200,000 in a county of about 300,000. A cluster of government buildings—including the state capitol, federal and state courthouses, city and county government offices, and the Madison Police Department (MPD)—are clustered together in the heart of downtown. The capitol is surrounded by Capitol Square, with park-like grounds and wide sidewalks. Connecting the capitol area with the University of Wisconsin campus is State Street, a half-mile pedestrian mall lined with stores and restaurants. The campus end of State Street creates an ell with Library Mall, a plaza flanked by campus libraries. The university's Bascom Hill rises just beyond Library Mall. There are small concrete elevated podiums and stage areas at each end of State Street that may be used by anyone who wishes to give a performance or speech; a permit to plug into the city's electricity at either site costs $5.25. Although a shortage of parking and the visible presence of deinstitutionalized mentally ill and homeless persons divert much of Madison's retail trade away from downtown, government employees and university students still dominate downtown spaces, and the downtown area remains the center of public life in the city. Downtown public life is seasonal. When it is warm enough, licensed food carts sell a variety of ethnic foods at the two ends of the mall, and performers and speakers use the two public stages throughout the day. Public ceremonies are most commonly conducted within the capitol rotunda or on the steps of the capitol, which makes an attractive backdrop. The capitol area is the usual site of protests and demonstrations,

although student rallies may remain in the Library Mall area or on Bascom Hill. Marches most often run between the Library Mall area and the capitol, or around Capitol Square. This traditional protest area intersects the jurisdiction of three separate police agencies. The Capitol Police have jurisdiction over the capitol and its grounds, the University of Wisconsin campus police cover all university areas (including Library Mall), and the Madison Police Department has jurisdiction over other downtown areas, including Capitol Square across the street from the capitol grounds, all of State Street, and the rest of the city.

This study focuses on 1994 and examines any public event for which we could find a record in a wide variety of official sources. This particular year was chosen because it was the most recent complete year when data collection began and was the year for which we had the most complete police records, since some agency records had been discarded for prior years. Other analyses from this project examine a more limited set of events and official sources over a longer time period (Oliver and Maney 1998).

Public Agency Sources

Each police agency in Madison has its own geographic jurisdiction and internal record-keeping logic. Most of the records were kept unsystematically, and the amount of information available varied markedly between sources. Records of permit applications were generally more systematic and contained more information than log book records. Log books provided information about unpermitted events, which frequently lacked details about the numbers, actions, identities, or issues of protesters. Although some researchers have treated a particular systematic police record as if it were a meaningful universe of events (e.g., Hocke 1998; McCarthy, McPhail, and Smith 1996), this apparent comprehensiveness is illusory. Police agencies have different physical and legal jurisdictions that lead them to record some events and not others and to be more comprehensive about some locations than others. A separate methodological study (Maney and Oliver 1998) explores these patterns for Madison and finds that events recorded by newspapers but not police largely occur indoors on private property (and thus outside police jurisdiction) or are small and nondisruptive and concludes that all record sources must be treated as incomplete. Different record sources must be assessed against each other to determine their logic of inclusion and exclusion of events. Thus, for each agency, we sketch the logic whereby it creates records.

*MPD parade permits.*—Parade permits are required for any event that will disrupt traffic on a public street, including road races, charity walks, and children's bike parades. These are considered temporary records and

archives for prior years are not available. We coded all MPD parade permits.

*MPD log.*—All 130,000 entries in the paper copy of the MPD 911 log book were scanned visually by the second author for items of potential interest. Although MPD log records are computerized, there is no Uniform Crime Report (UCR) category for public events and no search logic for retrieving them from computer files. A dozen unpermitted protests were located from the visual scan of the log; these records were created because of citizen complaints. Additionally, we identified several dozen incidents of collective violence (fights involving 10 or more people), which are excluded from the present analysis.[2] We found virtually no overlap between the log book and the permit records. MPD officers stated that it was their policy to create no record for a law-abiding event (permitted or unpermitted), even if officers spent significant amounts of time policing or observing it. It is legal to march or gather on city sidewalks without a permit as long as one does not impede the flow of traffic.

*Street-use committee.*—Street-use permits are required if a street is to be closed for more than a few minutes. These events are coordinated through an interagency street-use committee. These are considered temporary records and are unavailable for prior years. We coded all paper permit applications and committee agendas for 1994. The director of the street-use committee has also maintained a list of downtown events as handwritten notes in planning diaries, which we read for information about additional events. These records overlapped somewhat with other permit records, but also included events not recorded elsewhere.

*Capitol police permits.*—The Capitol Police are responsible for all state-owned property in Wisconsin. Permits are required for any gathering, display, or literature distribution on state property, and the data are recorded in a standardized format that makes this a comprehensive record of permitted events on state property. We were able to obtain a computer download of these permit records. Weddings, private meetings (e.g., for employees regarding their benefit options), and photography sessions were excluded from analysis.

*Capitol police log.*—Information from the computerized Capitol Police log was obtained with the assistance of the police, who conducted searches and generated lists of candidate events. Preliminary work revealed that

[2] It is questionable whether these "big fights" should be included as public events. Of the 76 large fights located in police records, we could find media traces of only two. We have not yet conducted exhaustive systematic searches for all traces of these fights, and it is possible that a couple more might have received some media mention. However, most of these fights resulted in no arrests, and many dispersed as soon as (or even before) the police showed up. None had any indication of having a "message." We plan a separate study of the phenomenon of collective violence.

the relevant event codes such as "crowd management" and "riot/rally" did not capture all relevant stories, so these were supplemented with keyword searches. For each candidate event, we examined and coded all relevant information in both the full log file and, when available, the computerized officer's report file. Similar to the MPD log, the Capitol Police log contains no systematic information about permitted events, but it provided often cryptic clues about unpermitted protest events. Small and uneventful protests generate log entries that only indicate a protest occurred with little or no information about the message, behavior, or number of protesters. Even more complete records of disruptive protests, indicated by phrases like "protesters climbing walls" or "people in wheelchairs heading for the governor's office," often lacked information on the identity of the protesters or the issues involved.

*Campus police log.*—With the aid of police staff, we obtained a computer download of potentially relevant campus police log book records that included up to five lines of "disposition" text. Paper and computerized report files were located and read for further information about events deemed relevant from the disposition. The campus police make no systematic attempt to record information on peaceful permitted events. Groups that wish to use university property (outdoors as well as indoors) are required to obtain the equivalent of a permit. However, these requests are handled through the university's room scheduling office, which logs literally thousands of space reservation records a month and has no search logic that would permit the location of the events of interest.

Media Sources: Local Newspapers

Madison has two daily newspapers, both listed as "midwest regional sources" by NEXIS. The *Capital Times* is a locally owned afternoon paper that does not publish on Sunday and that circulates principally in the Madison area. The morning *Wisconsin State Journal* is owned by Lee Enterprises, has about three times the circulation of its competitor, and is distributed more broadly across southern Wisconsin. The papers share production facilities (which are owned by a jointly owned holding company), but were founded separately and have distinct editorial policies and reporting staff. Editorially, the *Wisconsin State Journal* defines itself as moderate and politically independent, endorsing both moderate Republicans and moderate Democrats. The *Capital Times* defines itself as progressive and liberal Democratic. Rigorous computerized searches for events were conducted with the NEXIS database using all descriptors appearing in the police record as keywords, including actions, locations, participating individuals or groups, and synonyms for these. Every article

that explicitly mentioned the event was saved and coded, regardless of its length, location in the newspaper, or detail in describing the event.[3]

*Specifying the dependent variable (timely coverage).*—In the present analysis, we focus on "timely coverage," which we define as at least one unambiguous reference to the event in either newspaper during the 31-day interval starting 15 days before the event to 15 days after. Substantively, receiving timely coverage is necessary if an event is to have an impact through media coverage. The choice of this particular interval is based on preliminary analysis in this project using a broader 18-month interval (six months before to 12 months after events) to find all possible stories for message events and a subset of events not explicitly related to a message. Contrary to common assumptions in event research,[4] news stories about collective events are often printed many days or even weeks and months after the event and, in addition, prior coverage of events is quite common and may contribute to mobilization. If one wants to examine the total *volume* of coverage of an event, it is necessary to examine a broader interval around the event, or many stories will be missed.

However, although references to events can be found months before and after events, especially as the anniversary of an event approaches, the distribution of coverage is highly spiked in the several days before and after the event. Our preliminary analysis indicates that although only about half of the events receive "next day" mention, depending on the year and the types of events, 68%–78% receive mention within one day before or after the event, from 80% to 85% are mentioned within the week of the event (i.e., less than four days before or after), and 87%–90% are mentioned within the 15 days before or after the event. Outside the spike, the distribution of event coverage is relatively flat throughout the rest of the 18-month interval with the exception of a slight revival as the anniversary of the event approaches. A few events are mentioned *only* on their anniversaries, without having received coverage at the time of the event. However, references to events that receive only nontimely coverage are usually very short and incomplete. In sum, if one is interested only in whether an event receives some media coverage or not, media searches must not be too narrowly confined to the day of the event, but expanding the search more than a couple of weeks around the event provides very little payoff for the effort expended.

[3] Subsequent analyses will consider the content and framing of news coverage. The first step is simply to determine the factors that lead an event to receive any mention at all.

[4] European studies of protest events have tended to focus on the Monday issue of newspapers and to restrict their attention only to events occurring in the previous weekend (e.g., Kriesi et al. 1995; Rucht and Neidhardt 1998).

Of the 121 events that received any newspaper coverage, 82% (99) were covered by both newspapers, 4% (5) were covered only by the *Wisconsin State Journal,* and 14% (17) only by the *Capital Times.* Although the selection logics of the two newspapers were very similar, there were a few crucial differences between them, so separate results are presented for each.

## RESULTS

### The Shape and Rhythm of Public Life

When activists stage protests or demonstrations to call public attention to their claims and concerns, they do so against the backdrop of public life in public spaces. What else is going on in public space? Table 1 shows the distribution of event types in the official records and the percentage of each type that received timely media coverage. The "proportion of records" column includes all relevant records, while the rest of the table excludes the 127 literature distribution permit records.

Fully 25% of the police records are Capitol Police permits for literature distribution. Even this percentage is understated as a percentage of events, because most literature distribution permits are for multiple dates, usually many or all of the Saturdays between April and October, at the Farmer's Market on Capitol Square. As the first set of columns in table 2 indicates, 34% of these permits were to social movement organizations, and another 34% were to issue-oriented organizations, thus indicating that a wide variety of organizations were seeking to influence public opinion in this way. No literature table received newspaper mention, which is not particularly surprising, so we dropped literature distribution from the analysis of the correlates of media coverage.[5]

Standard protest forms that have been the objects of "protest events" research (rallies, marches, vigils, and unpermitted protests) are only a

---

[5] McCarthy, McPhail, and Smith (1996) keep literature tables and leafleting in their models despite the low frequency of press coverage of such events. We did not because, following our exhaustive event-specific protocol, doing so would require creating a separate event for each distinct date and then concluding what we already know, which is that these activities did not receive press coverage. Although it is possible that a more systematic and exhaustive media search might uncover an occasional mention of a literature table, the cost/payoff ratio from such an enterprise would be prohibitive. Similarly, we excluded from the analysis the Farmer's Market itself, which occurs every Saturday in season and receives a great deal of media attention, both in periodic articles about the Farmer's Market and in passing mention to incidents which occur "at the Farmer's Market." Trying to determine whether every distinct Saturday's market did or did not receive its own particular mention was deemed a pointless enterprise.

TABLE 1

FREQUENCY DISTRIBUTION AND MEDIA COVERAGE OF EVENT TYPES

| Event Type | N | Proportion of Records | Proportion of Events* | Any Media Coverage (%) | *Capital Times* (%) | *Wisconsin State Journal* (%) |
|---|---|---|---|---|---|---|
| Standard protest events .......... | 49 | 09 | .13 | 44 | 42 | 33 |
| Rally† .. ... .. .... ............... .. .. | 9 | .02 | 02 | 78 | 78 | 56 |
| March or moving rally ....... ... | 11 | .02 | 03 | 54 | 54 | 45 |
| Vigil ................... .. ... . ...... | 5 | 01 | 01 | 20 | 20 | 0 |
| Protest, unpermitted‡ . . . .. | 24 | .05 | .06 | 33 | 29 | 25 |
| Other message events ............. . | 105 | .21 | .27 | 51 | 50 | 46 |
| Ceremony§ .. ... ... ............... | 34 | .07 | .09 | 59 | 56 | 47 |
| Speech, hearing ... ....... . ......... | 11 | .02 | 03 | 45 | 45 | 36 |
| Mixedᴵ ....... .. ... .. ................. | 45 | 09 | 12 | 40 | 40 | 38 |
| Fair, sale, commercial• . .. | 15 | .03 | .04 | 73 | 73 | 73 |
| Message nonevents ...... .. ... ... .. | 169 | .33 | .11 | 26 | 24 | 19 |
| Literature table .................... | 127 | .25 | . . . | | | |
| Display** ....... ... ... .. .. ... ..... | 29 | 06 | .08 | 24 | 21 | 14 |
| Collection, distribution†† .... . | 13 | 03 | .03 | 31 | 31 | 31 |
| Other events . ... ................. .. | 186 | .37 | 49 | 19 | 18 | 17 |
| Public social event‡‡ ............. | 60 | .12 | 16 | 12 | 12 | 12 |
| Athletic event§§ ... ... .. ......... | 26 | 05 | .07 | 50 | 50 | 42 |
| Concert or performanceᴵᴵ .. | 91 | 18 | .24 | 11 | 9 | 11 |
| Parade . . .. ................... . .. | 9 | .02 | .02 | 44 | 44 | 44 |
| All records .. ... .. .. ................. | 509 | · | | 32 | 30 | 27 |

* *N* for proportion of events = 382, literature tables are excluded
† Includes one combination of a lobbying day and a rally
‡ Includes unpermitted literature distributions that resulted in police complaints
§ Includes ceremonies, memorial services, and ceremony/display events, includes three press conferences.
ᴵ Includes social, entertainment, or athletic events with clear informational or fund-raising purposes indicated in the official record; includes Take Your Daughter to Work day
• Includes five commercial promotions which all received media coverage.
** Includes both informational and art displays, includes Christmas light tours
†† Includes collections of items other than money (food, clothing, blood) and distributions of items other than literature (poppies and forget-me-nots for Memorial Day and Veterans Day)
‡‡ Includes student, church, neighborhood parties, and neighborhood children's parades as well as larger public recreational events that are not athletic and do not include performances
§§ Includes competitive and fun walks, runs, bike rides with no mention of fund-raising or message; includes a vintage car race.
ᴵᴵ Includes a big fireworks show

small proportion of all events occurring in public spaces; they are only 9% of all records, and only 13% of event records excluding literature distribution. However, some of these events that took on protest *forms* involved nonprotest *issues,* including marches against drugs, against suicide, in favor of church attendance, and promoting a multiethnic festival, along with a vigil commemorating the Holocaust.

## TABLE 2

TYPES OF SPONSORING ORGANIZATIONS FOR EVENTS AND LITERATURE TABLES

| TYPE OF ORGANIZATION | LITERATURE TABLES N | LITERATURE TABLES Proportion | EVENTS N | EVENTS Proportion |
|---|---|---|---|---|
| None* ... .. ............. ... ................... ............. .. ... ...... | | | 96 | .25 |
| Nothing in record ............. .. ................. .. ........ | 0 | | 69 | .18 |
| City events office personnel ....... .. ..... ..... ...... | 0 | | 27 | .07 |
| Event specific* ....... . ... ............. .. ... ................. .. | | | 29 | 08 |
| Event-specific organization ..... ... .. ................ | 3 | .02 | 24 | .06 |
| Individual name only ............... ..... .................... | 1 | .01 | 5 | .01 |
| Government* ..... .. .. ....................... ............. .. ..... | | | 32 | .08 |
| Government agency ................. ............ .......... | 4 | .03 | 26 | .07 |
| Elected official, candidate, or party ............... | 8 | .06 | 6 | .02 |
| Occupational* ............... ............... .. ................. . | | | 9 | .02 |
| Union or labor group ............... ... ............. ... | 0 | | 3 | 01 |
| Professional association . ................. .. ............. | 0 | | 6 | .02 |
| Public service* ............ ...................... .. .. ........... | | | 10 | .03 |
| Military, veterans group .......... .. .. ............. ... .. | 0 | | 6 | .02 |
| Service club ........ . ...................... .. ............ ... ... .. | 0 | | 4 | .01 |
| Social movement organization: | | | | |
| Local* ..... ............ ...................... .................... | 21 | .17 | 6 | .02 |
| National* .................... .. ................ . ..... .......... | 9 | 07 | 10 | .03 |
| Issue oriented:† | | | | |
| Local* .. ............. ............... .. ... .. ............. .... | 22 | .17 | 17 | .04 |
| National* ......... .. ............... ... . ............. . ... . | 9 | .07 | 25 | .07 |
| Other: | | | | |
| Religious group* ....... .. .. ............... . .. .. ............ | 8 | .06 | 16 | .04 |
| Business association* ............... .. ................... .... | 5 | .04 | 10 | .03 |
| Business (particular)* ... ................... ... . ....... | 1 | 01 | 17 | .04 |
| University* ....................................... .. .. ........... | 0 | | 8 | .02 |
| Nonprofit institution* ..... ...... .......... . .. ........ | 17 | .13 | 11 | .03 |
| University student group* ...... ................ ........ | 1 | .01 | 18 | .05 |
| Youth group or school* ....... ............. .. ........... | 3 | 02 | 23 | .06 |
| Recreational group* ............... ... .. ............ ... .. . | 10 | .08 | 30 | .08 |
| Neighborhood association* ................ ............ | 1 | .01 | 15 | .04 |
| Total ................. ... ............... .. ....... .. .. ..... | 127 | 1.00 | 382 | 1.00 |

* These organization types are used as independent variables in multivariate models; infrequent types with conceptual similarities and similar rates of media coverage are grouped to avoid estimation problems
† Educational, charitable, or advocacy groups

With literature tables excluded, the purpose of over half the events is solely entertainment or recreation. The largest categories are performances and parties. There were 60 permits granted to close streets for public parties (most of these were for neighborhood parties or children's parades) and another 90 permits were granted for individuals or groups to perform on public property, generally inside the capitol building or in the Capitol Square area. Even this figure for performances is underestimated, for the street-use coordinator's planner contained another 125 records that appeared to be performances by individuals and groups in the downtown area. We did not systematically code or search media sources for these additional performance events.

In addition to these entertainment events, our analysis of the official records revealed a large number of events that were clearly vehicles to persuade or inform the public even though they were not protest events. Most important among these for scholars of social movements are the ceremonies and the events that mix messages with recreation. A ceremony is a nondisruptive permitted event in a public place at which people give an award, memorialize someone, dedicate something, or give a short speech in honor of or expressing concern about something. These events often appeared to be staged "media events." The three press conferences were grouped with them because of their very similar form.[6] We coded eight ceremonies as not having a "message" designed to influence the public.[7] Most of the mixed events were fund-raisers, that is, social, recreational, or entertainment events that raise money for a particular organization or cause. We also included in this group a few additional social and recreational events that did not mention fund-raising, but were clearly intended to convey a message related to public issues (e.g., the NAACP's Freedom Fest or the "fun walk in the watershed," whose stated purpose was to inform the public about watershed issues).

The last three columns of table 1 show the proportion of each event type that received media coverage for both papers combined and for each paper separately. Notably, the highest rate of any media coverage is for rallies, followed by commercial events and then ceremonies. In the next tier are athletic events, marches, speeches and hearings, and parades. It

[6] It might be argued that press conferences should not be included, as all three obtained media coverage. However, the media do not always show up when someone invites them to a press conference, and, in their form, press conferences are very similar to many of the "ceremonies." If press conferences are excluded, 71% of the ceremonies received media coverage instead of the present figure of 74%.

[7] These were military and student award presentations, a building dedication, a meeting with the press that appeared to be about security procedures, ceremonies for art displays, and an athletic event. All were ambiguous cases in which a message might have been intended; 60% of these events received media coverage.
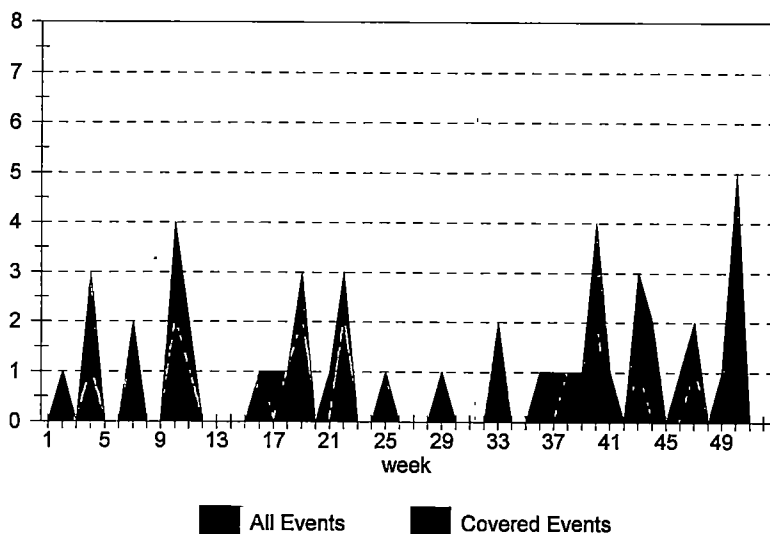
FIG. 1.—Weekly frequency of events: protests and demonstrations

is also worth noting that the "message nonevents"—displays, distributions, and collections—receive more media coverage on average than the amateur performances or public social events, which are much larger and involve some actual activity. The two newspapers are broadly similar in their patterns of event coverage. However, consistent with the difference in editorial policies, the Capital Times is considerably more likely to cover rallies and somewhat more likely to cover other protest forms and speeches and ceremonies than the Wisconsin State Journal.[8]

Cyclical and noncyclical variations play an important role in public events. Nearly a third of all events occur on Saturday. Figures 1–4 show the weekly frequencies of different groups of events throughout the year, demonstrating both clear seasonal cycles and wide weekly fluctuations within those cycles. There are few events early in the year in the depths of winter. All activity drops off during the vacation period of spring break (week 13). Social and recreational events mostly occur in the warmer weather between weeks 16 (mid-April) and 44 (end of October), with a peak on July Fourth (week 27). Protests and demonstrations and other message events are less common during the summer vacation months (weeks 23–35); the "other" message events tend to occur mostly in the

---

[8] This difference is larger in 1994 than in the other three years (1993, 1995, 1996) for which we have data.
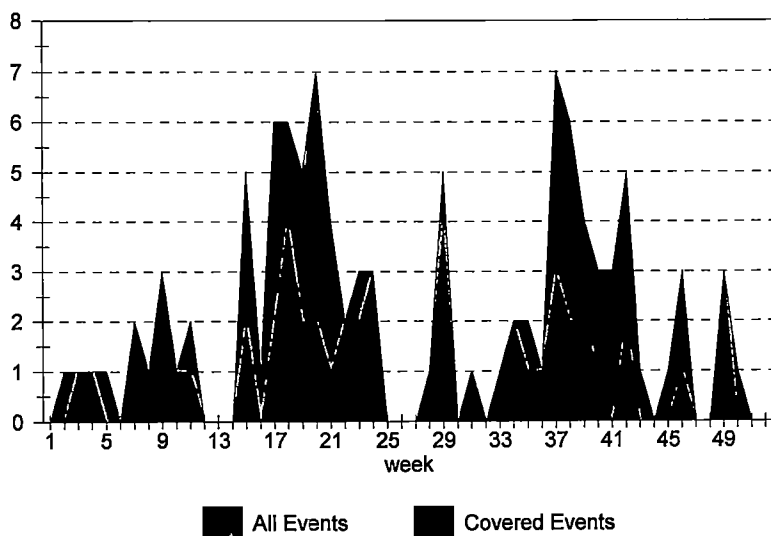
FIG. 2.—Weekly frequency of events: other message events
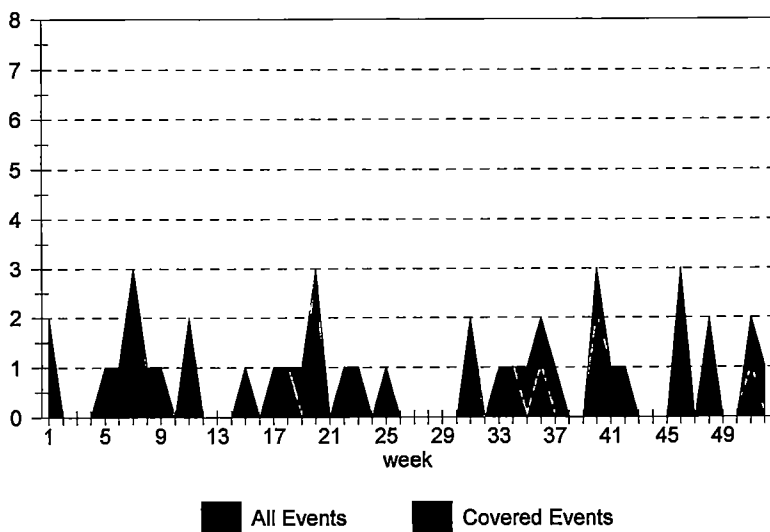


FIG. 3.—Weekly frequency of events: displays and collections
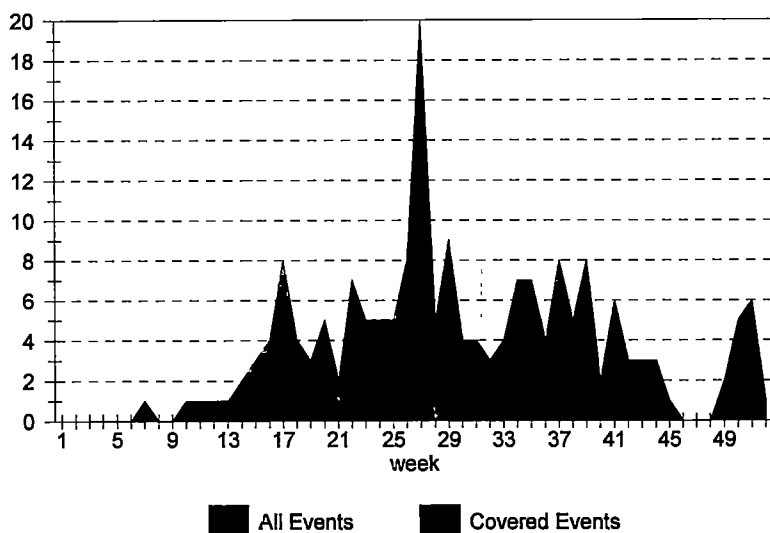
FIG. 4.—Weekly frequency of events: entertainment events

spring and fall, while protests and demonstrations are more scattered through the school year. Collections and distributions are more evenly distributed through the year. The seasonal pattern is, not surprisingly, most pronounced when indoor and outdoor events are distinguished. As figure 5 indicates, the seasonal patterns are somewhat complementary, with indoor events peaking in spring and fall and dying out in the summer, while outdoor events peak in the summer. However, both kinds of events are rare in January and February, and both kinds of events rise in the spring and fall.

Examination of the superimposed plots of coverage rates for the event types in figures 1–4 suggests that events have a higher probability of coverage if they occur when fewer other events are occurring in the same week. This possibility is examined below.

### Correlates of Media Coverage

In this section, we describe the independent variables we coded. We also present their bivariate relations with media coverage.

### Size

It is well established that the size of an event is a major predictor of its media coverage (McCarthy, McPhail, and Smith 1996; McCarthy et al.
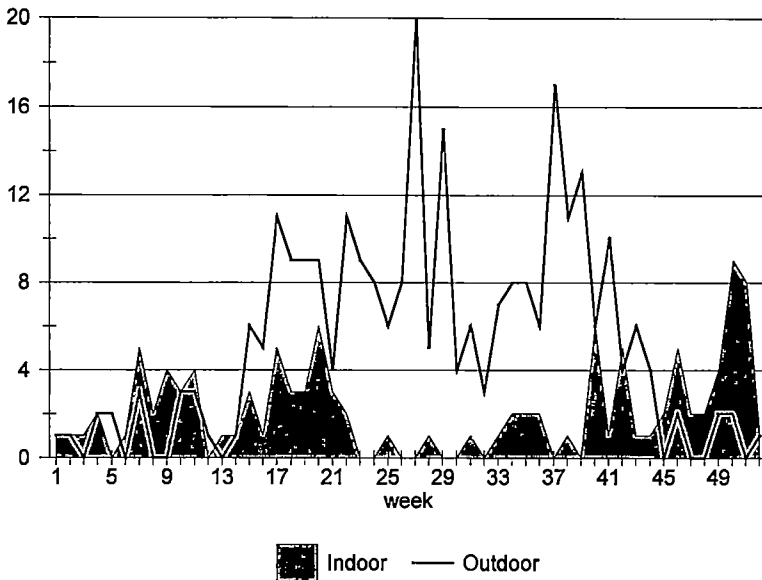
FIG. 5.—Seasonal pattern of indoor and outdoor events

1996; Snyder and Kelly 1977). Unfortunately, numerical estimates of the number of participants are missing from the large majority of Madison police records, and certain event types are more likely to have size estimates than others. We looked for size-relevant information in all records, including equipment requests for chairs. Even with these kinds of clues, 59% of the police records lack any size information. So in addition to using the objectively available information on event sizes, we coded a subjective size variable on the basis of whatever information and impressions we could glean from police comments, the sizes of other similar events, or our own knowledge of local events. This subjective size variable is obviously imperfect, but adequately captures the differences among small, medium, large, and very large events. After preliminary analysis using a variety of coding schemes (detailed in table 3), the size variable used in the analysis is the four-category scheme shown in table 3 because it has high predictive value, produces less distortion in the effects of other factors, and is closer to the gross distinctions of the original coding scheme. As table 3 indicates, the size of an event has a strong positive effect on the coverage of events. However, the smallest events, the nonevent displays, distributions, and collections, which were often coded as involving no people at all (i.e., having zero size) were almost as likely to be covered as medium-size events, while the small events (involving fewer than 16 people) were much less

## TABLE 3

### CODING OF SIZE OF EVENTS AND NEWS COVERAGE BY EVENT SIZE

| | ESTIMATED SIZE | | | | SIZE IN RECORD* | | | COMBINED SIZE GROUPS | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CODING | Coded Categories (N Persons) | N | Midpoint | Log Midpoint | Mean* | N* | Any News (%) | GROUP LABEL | N | Any News (%) | CT (%) | WSJ (%) |
| 0 ......... | None† | 39 | 0 | 0.0‡ | | 0 | 26 | ... | 39 | 26 | 23 | 18 |
| 1 ......... | Tiny (1–5) | 19 | 3 | 1.1 | 3 | 7 | 5 | Small | 54 | 7 | 7 | 7 |
| 2 .. ...... | Very small (6–15) | 35 | 10.5 | 2.4 | 12 | 8 | 9 | | | | | |
| 3 ......... | Small (16–30) | 62 | 23 | 3.1 | 29 | 12 | 29 | Medium | 234 | 28 | 27 | 23 |
| 4 ... ...... | Modest (31–99) | 86 | 65 | 4.2 | 56 | 38 | 28 | | | | | |
| 5 ... ....... | Medium (100–499) | 86 | 300 | 5.7 | 203 | 56 | 28 | | | | | |
| 6 ... ........ | Larger (500–1500) | 25 | 1,000 | 6.9 | 695 | 16 | 72 | Large | 55 | 75 | 73 | 71 |
| 7 ... .. ...... | Large (2,000–10,000) | 25 | 6,000 | 8.7 | 4,375 | 8 | 72 | | | | | |
| 8 ............ | Very large (10,000+) | 3 | 45,000 | 10.7 | | 0 | 100 | | | | | |
| 9 ............ | Huge (100,000+) | 2 | 200,000 | 12.2 | 100,000 | 1 | 100 | | | | | |

NOTE.—CT = *Capital Times; WSJ = Wisconsin State Journal*

\* Observed size information uses any clues available in the record and is not based on consistent information.

† Displays, collections

‡ The log of zero is undefined, so the "log midpoint" variable is set to zero (i.e., the log of 1.0) for events of size zero, as this makes the scale behave most reasonably.

TABLE 4

AVERAGE SIZE OF EVENT TYPES

| TYPE | MEANS | | MEDIAN | | N | |
|---|---|---|---|---|---|---|
| | Observed | Estimated | Observed | Estimated | Observed | Estimated |
| Rally ... . ...... .. .. ...... | 265 | 181 | 265 | .65 | 2 | 9 |
| March ..................................... | 940 | 766 | 250 | 300 | 5 | 11 |
| Vigil ............................. .. ... . | | 23 | | 23 | 0 | 5 |
| Ceremony ............................. | 76 | 86 | 100 | 23 | 11 | 34 |
| Speech ...... ... .. .. .. .. ...... .. | 75 | 661 | 75 | 65 | 2 | 11 |
| Protest ........... .. . .. ... ...... . | 20 | 72 | 11 | 11 | 16 | 24 |
| Display ..................... . .. .... . | 0 | 10 | 0 | 0 | 1 | 29 |
| Distribution, collection ...... | 100 | 46 | 100 | 0 | 2 | 13 |
| Mixed ............ .................... | 3,942 | 5,369 | 300 | 300 | 31 | 45 |
| Commercial .. .... .... .. ......... . | 300 | 11,566 | 300 | 6,000 | 1 | 15 |
| Social .......... ....... .. .... . ... | 160 | 381 | 62 | 65 | 50 | 60 |
| Athletic ................................. | 958 | 1,904 | 500 | 300 | 11 | 26 |
| Performance ...... ... .............. | 66 | 2,341 | 11 | 23 | 22 | 91 |
| Parade . ...... . .. ... ... ... ... .. | 1,808 | 2,200 | 150 | 300 | 6 | 9 |

NOTE —Observed size statistics are based on those events for which some size information was available in the police record Estimated size statistics use the midpoint of the subjective size category, as shown in table 3

likely to be covered, and the largest events (involving 500 or more people) were much more likely to be covered.

However, as table 4 shows, differences in rates of media coverage between types are not just due to the sizes of events. "Mixed" events are considerably larger than ceremonies and larger than rallies and marches but are less likely to receive media coverage. This suggests the importance of more detailed information about the factors affecting news coverage.

*Organizations and Media Access*

To assess the effects of organizational sponsorship on media access, we developed the detailed classification shown in table 2 of the types of organizations listed in official records as sponsoring or connected with the events. Organizations classified as "national" are the local chapters of national organizations, and are contrasted with purely local groups. Social movement organizations (SMOs) are voluntary associations whose purpose is the pursuit of social change goals in a contentious fashion. By contrast, the issue-oriented groups are voluntary associations that pursue social issues and concerns in a more consensual fashion, such as the American Heart Association or the Children's Trust Fund. The issue-oriented

groups would generally see themselves and be seen by the movement groups as being very different from them, although they typically mix advocacy with fund-raising and service. The "nonprofit" category includes museums, orchestras, hospitals, social service agencies, and development groups that are professionalized and that receive significant government funding. "Business associations," which promote an industry or business sector, are distinguished from particular business firms. Businesses are not necessarily promoting their own ends when they sponsor events: some are for-profit firms whose business is to put on public celebrations, while in other cases businesses sponsor events as a kind of public service.

Comparing the organizations represented in literature distribution permits to those sponsoring other event forms in table 2, it is clear that many organizations seeking to communicate messages to the public do so directly, with literature distribution. Movement organizations, issue organizations, religious groups, and nonprofits dominate literature distribution. The pool of organizations distributing literature might be viewed as a kind of inventory of organizations or messages seeking public attention.

Table 5 shows the rate of newspaper coverage for events according to organizational sponsorship. The two newspapers are generally quite similar in their responsiveness to the various organization types. The major exception concerns the national SMOs: consistent with editorial policies, the *Capital Times* covered 70% of their events (nearly tying its coverage of business association events) while the *Wisconsin State Journal* covered only 40%. Business associations have by far the highest rate of coverage in both newspapers, with 90% of their events covered. About half the events of other "insider" groups are covered by both newspapers, including military and veterans groups and service clubs, particular businesses, the university, and nonprofits. It is notable that event-specific organizations also have a relatively high success rate of 59%. These are often established organizations that put on the same large event year after year. Lower rates of coverage, around one-third, are obtained by occupational groups, local SMOs, and the consensual issue-oriented health, education, and charitable organizations, both national and local. Events sponsored by recreational groups and neighborhood associations have even lower rates of coverage. Finally, events sponsored by religious groups or youth groups or schools are less likely to receive coverage than events with no organization mentioned at all.

*Location*

Although it is well recognized that news routines, "beats," and the physical constraints of preparing news stories for a deadline are important factors in determining whether an event is covered, prior researchers have not

TABLE 5

MEDIA COVERAGE OF EVENTS SPONSORED BY DIFFERENT TYPES OF ORGANIZATIONS

| Type of Organization | N | Any News (%) | CT (%) | WSJ (%) |
|---|---|---|---|---|
| None* | 96 | 14 | 12 | 11 |
| Nothing in record | 69 | 19 | 16 | 14 |
| City events office personnel | 27 | 4 | 4 | 4 |
| Event specific* | 29 | 59 | 59 | 59 |
| Event or campaign specific | 24 | 58 | 58 | 58 |
| Individual name only | 5 | 60 | 60 | 60 |
| Government* | 32 | 31 | 31 | 25 |
| Government agency | 26 | 31 | 31 | 23 |
| Elected official, candidate, or party | 6 | 33 | 33 | 33 |
| Occupational* | 9 | 33 | 33 | 11 |
| Union or labor group | 3 | 33 | 33 | 0 |
| Professional association | 6 | 33 | 33 | 17 |
| Public service* | 10 | 50 | 40 | 50 |
| Military, veterans group | 6 | 50 | 33 | 50 |
| Service club | 4 | 50 | 50 | 50 |
| Social movement organization. | | | | |
| Local* | 6 | 33 | 33 | 33 |
| National* | 10 | 80 | 70 | 40 |
| Issue oriented: | | | | |
| Local* | 17 | 35 | 35 | 35 |
| National* | 25 | 36 | 32 | 32 |
| Other: | | | | |
| Religious group or congregation* | 16 | 12 | 12 | 12 |
| Business association* | 10 | 90 | 90 | 80 |
| Business (particular)* | 17 | 47 | 47 | 41 |
| University* | 8 | 50 | 50 | 50 |
| Nonprofit institution* | 11 | 45 | 45 | 45 |
| University student group* | 18 | 39 | 39 | 39 |
| Youth group or school* | 23 | 4 | 4 | 4 |
| Recreational group* | 30 | 27 | 27 | 17 |
| Neighborhood association* | 15 | 20 | 20 | 20 |
| All organization types | 382 | 32 | 30 | 27 |

* These organization types are used as independent variables in multivariate models; infrequent types with conceptual similarities and similar rates of media coverage are grouped to avoid estimation problems.

drawn the logical inference that the spatial location of an event would be related to reporters' routines. The "state government" beat is physically located at the capitol, for example, while the "university" beat is located at the University of Wisconsin, Madison. Location information in the police records was carefully coded into a detailed set of categories capturing the exact location of all events. Based on both the logical characteristics of

TABLE 6

CODING OF LOCATIONS AND NEWS COVERAGE BY LOCATION

| | N | PROPORTION | MEDIA COVERAGE (%) | | |
|---|---|---|---|---|---|
| | | | Any | CT | WSJ |
| Downtown/university inside: | | | | | |
| Inside capitol | 66 | .17 | 36 | 33 | 32 |
| Inside other downtown building | 12 | .03 | 42 | 33 | 33 |
| Inside university building | 13 | .04 | 54 | 46 | 54 |
| Downtown/university outside: | | | | | |
| Capitol steps, grounds | 79 | .21 | 29 | 28 | 25 |
| Other downtown outside | 49 | .13 | 57 | 57 | 43 |
| University outside except Library Mall | 12 | .03 | 58 | 58 | 50 |
| Library Mall area | 29 | .08 | 7 | 7 | 7 |
| Not downtown: | | | | | |
| Inside | 19 | .05 | 21 | 21 | 16 |
| Outside | 103 | .27 | 20 | 20 | 19 |

spaces and the frequencies of media coverage of events in those spaces, these were distilled into the categories shown in table 6. Except for the Library Mall area, events in the downtown or campus areas receive more media coverage than those away from downtown. There appears to be no particular difference in the coverage of indoor versus outdoor events. The relatively low gross rate of media coverage for events inside the capitol or other downtown buildings and on the capitol grounds compared to other downtown areas is due to the mix of events in these spaces: most displays are in the capitol or other downtown buildings and most performances are on the capitol grounds or in the capitol rotunda. It is not immediately obvious why events on the Library Mall have such low rates of media coverage, although it may be due to the mix of events in the space or a lack of news interest in student-oriented events.

### Newsworthiness

*Messages and conflict.*—The first part of table 7 shows the bivariate relation between newspaper coverage and newsworthiness factors. If we understand newsworthiness as being about the public sphere, as about communicating information relevant to public concerns, then it is reasonable to suppose that events attempting to convey a message to a larger public ought to be more newsworthy than nonmessage events. We coded an

TABLE 7

MEDIA COVERAGE BY SELECTED CHARACTERISTICS (EVENTS ONLY)

| | | | MEDIA COVERAGE (%) | | |
|---|---|---|---|---|---|
| | $N$ | PROPORTION | Any | CT | WSJ |
| Newsworthiness: | | | | | |
| No message | 209 | .55 | 24 | 23 | 22 |
| Message | 173 | 45 | 41 | 39 | 34 |
| No conflict | 330 | .86 | 28 | 28 | 25 |
| Conflict | 52 | .14 | 52 | 48 | 40 |
| No message, no conflict | 208 | .55 | 24 | 23 | 21 |
| Consensual message | 122 | .32 | 37 | 36 | 32 |
| Conflictual message | 51 | .13 | 51 | 47 | 39 |
| No disorder | 358 | .94 | 31 | 30 | 27 |
| Disorder | 24 | .06 | 46 | 42 | 38 |
| No vehicles | 352 | .92 | 29 | 28 | 25 |
| Vehicles | 30 | .08 | 63 | 63 | 53 |
| No amplifier | 316 | .83 | 30 | 29 | 26 |
| Amplifier | 66 | .17 | 39 | 38 | 33 |
| Organizer not local | 26 | .07 | 15 | 12 | 8 |
| Local organizer | 356 | .93 | 33 | 32 | 29 |
| $N$ of police (quantitative factor): | | | | | |
| 0 | 335 | .88 | 29 | 28 | 25 |
| 1 | 15 | .04 | 47 | 47 | 40 |
| 2 | 14 | .04 | 29 | 29 | 29 |
| 3 | 6 | .02 | 67 | 67 | 17 |
| 4 | 3 | .01 | 67 | 67 | 67 |
| 5 | 5 | .01 | 40 | 40 | 40 |
| 6 | 1 | .00 | 100 | 100 | 100 |
| 7 | 1 | .00 | 100 | 0 | 100 |
| 8 | 1 | .00 | 100 | 100 | 100 |
| 9 | 1 | .00 | 100 | 100 | 100 |
| Timing and routines: | | | | | |
| Not annual | 333 | .87 | 27 | 26 | 22 |
| Annual | 49 | .13 | 63 | 61 | 61 |
| Not holiday | 350 | .92 | 31 | 29 | 26 |
| July Fourth | 16 | .04 | 19 | 19 | 19 |
| Other holiday | 16 | .04 | 69 | 69 | 63 |
| Legislature in session | 53 | .14 | 30 | 31 | 28 |
| Legislature not in session | 329 | .86 | 32 | 28 | 25 |
| Indoors | 110 | .29 | 36 | 33 | 32 |
| Outdoors | 272 | .71 | 30 | 29 | 25 |
| Day of week: | | | | | |
| Sunday | 52 | .14 | 34 | 35 | 27 |
| Monday | 52 | .14 | 40 | 40 | 35 |
| Tuesday | 41 | .11 | 27 | 24 | 24 |
| Wednesday | 30 | .08 | 23 | 20 | 17 |
| Thursday | 28 | .07 | 32 | 29 | 25 |

TABLE 7 (*Continued*)

| | N | PROPORTION | MEDIA COVERAGE (%) | | |
|---|---|---|---|---|---|
| | | | Any | CT | WSJ |
| Friday ...................... ... ..... | 56 | .15 | 34 | 34 | 30 |
| Saturday ..... .. .. ... .. ................... .. . .. | 123 | 32 | 29 | 28 | 27 |
| Time of day: | | | | | |
| All day ......................... .. ... ...... ................ | 78 | 20 | 33 | 32 | 29 |
| Morning ..... . . ..  ... .... ................. ..... | 70 | .18 | 33 | 31 | 29 |
| Midday . .. .. . ...... ................ . . ... | 88 | .23 | 28 | 27 | 24 |
| Afternoon ......................... .............. . .. ........ | 52 | .14 | 37 | 35 | 31 |
| Evening .............. . .. ............................ | 94 | .25 | 30 | 29 | 26 |

event as having a "message" if part of its apparent purpose was to influence public opinion or action, rather than the event being an end in itself. Standard protest events (rallies, marches, vigils, unpermitted protests) as well as literature distribution, collections, and commercial promotional events were all assumed to have messages. Social or athletic events whose stated purpose included raising money or informing the public were classified as "mixed" events and deemed to have message content. Similarly, public social events with clear "ethnic" labels were considered to be conveying a message of ethnic pride. For other events, we made judgments based on the nature of the sponsoring group and the police description of the activity to determine whether a message was involved. Most speeches, ceremonies, and displays were judged to have message content, although a minority appeared to be ends in themselves (e.g., student awards ceremonies, art displays, historical lectures). Most performances and social events were judged to lack message content. As table 7 indicates, message events receive substantially more coverage than nonmessage events, although this effect is stronger for the *Capital Times*.

The element of drama or contention is recognized as making an event newsworthy. We defined an event as involving "conflict" if we were aware of social or political conflict about the sponsoring organization or the issue addressed by the event. One event, the Mifflin Street block party, was judged to involve conflict despite having no message, because of the history of the event and the substantial public debate about it. Otherwise, conflict events are a subset of message events. As table 7 indicates, conflictual events receive much more coverage than nonconflictual events.

Eliminating the Mifflin Street block party, we can trichotomize events as involving no message, a nonconflictual message, or a conflictual mes-

sage. Table 7 shows that the coverage of consensual-message events is midway between nonmessage and conflictual-message events.

*Other newsworthiness factors.*—Since the literature refers to an extremely wide range of factors as affecting newsworthiness and since our police records were generally sketchy and incomplete, we proceeded somewhat inductively and coded all information available in the records that might affect coverage. From log book records, we could code the number of different police officers mentioned in the record, which is generally an indicator of an unexpected disruption. Large permitted events generally yield no mention in log books, even when many officers are working overtime to police the event. However, the sudden emergence of several dozen unruly protesters often leads to entries in log books about officers being called in early or to reports from several different officers as events unfold. We also coded a more subjective assessment of whether the event involved disorder, which was based on comments in the police log about problems, including extensive trash, people upset about locked restrooms, and traffic problems, as well as more politicized disruptions. Permit records indicate whether there are vehicles involved and whether there will be electrical amplification of sound, features which might make an event seem more interesting or newsworthy. As table 7 indicates, all of these newsworthiness factors increase the likelihood of coverage. Events involving disorder, more than three policemen, vehicles, or an amplifier were more likely to be covered by both newspapers.

Finally, we examined the address of the contact person for permitted events and coded whether this person was from the Madison metropolitan area or elsewhere. Additionally, the descriptions of a few events in the log book make it clear when the participants are from out of town (e.g., when the protesters get off buses from Milwaukee). Even "national" newspapers emphasize local events and personalities, and the lifeblood of a local newspaper is the local angle. Additionally, local organizers are more likely to be in communication with local news reporters. As expected, events with nonlocal organizers were less likely to be covered.

*Timing and Competition for News Hole*

Newspapers publish daily and have rigid deadlines that create constraints on reporters to research and write their stories quickly. The "news hole" is the amount of space available for news stories in a particular issue of a particular publication: newspapers do not publish blank spaces and cannot readily add extra pages. When there is a lot of news, some gets left out. When news is "slow," editors will fill the pages with whatever they can find. Thus, the coverage of any particular news item is inevitably

affected by the presence or absence of competing news. News reportage of local events requires local news personnel and thus could be particularly susceptible to factors of timing and the competition between events.

The time of the event in relation to news deadlines is commonly cited as a news routine issue. Since the *Wisconsin State Journal* is a morning paper, while the *Capital Times* is published in the afternoon, the literature would suggest that timing would influence which paper covers an event. Permitted demonstrations at the capitol can occur only between 11 A.M. and 1 P.M., or between 4 and 6 P.M. Based on this rule and by inspecting the distribution of starting and ending times available in the records, event times were categorized as all day (beginning before 10 A.M. and ending after 2:30 but before 8 P.M.), morning (beginning before 10 A.M. and ending before 2:30 P.M.), midday (beginning after 10 A.M. and ending before 2:30 P.M.), afternoon (beginning after 10 A.M. and ending after 2:30 but before 8 P.M.), and evening (part of the event is after 8 P.M.). Table 7 indicates that the time of day makes little difference in an event's coverage by either newspaper.

Less has been said about day-of-the-week rhythms in coverage. Events occurring on different days of the week have different probabilities of coverage, with events occurring on Mondays being somewhat more likely to receive news coverage than events on other days of the week, and events on Wednesday least likely. It is not clear whether these daily differences arise from reporters' routines or from daily variations in the size of the news hole, but they appear similar for the two newspapers.

Since news coverage is daily and tends to be tightly compressed around an event, the number of competing events within a short time frame may influence the prospects of coverage, and it appears from figures 1–4 above that the probability of coverage is lower when there are competing events. To assess these effects, we calculated a variety of moving averages of the number of events around each date, and found that moving averages in the range of two weeks (between 11 and 19 days) have the strongest negative relation to media coverage in these data, with the 17-day average having the strongest bivariate negative correlation with media coverage. Substantively, this suggests that events compete for space in the newspaper with other events occurring in the preceding and following week.[9] As

[9] The effect of this variable is influenced by the "outlier" July Fourth, when there were 16 events; the next highest total is seven. The effect of the variable is weaker if July Fourth is removed from the data. A more detailed analysis of interactions also revealed that there is a strong negative correlation in 1994 between the number of message events in a day or three-day period and the probability of media coverage of a particular message event, while there is no such correlation for nonmessage events. However, a check of the 1993, 1995, and 1996 data for message events found no such relationship, so this is not reported or interpreted in this article.

table 7 shows, events are less likely to be covered when they are competing with more events in the 17-day period (eight days before to eight days after) around them.

Events that can be predicted in advance are often covered in different ways from breaking news. A dummy variable indicates whether the event is linked to a holiday or an anniversary (including Christmas-related events, New Year's Eve, Martin Luther King Day, Independence Day, St. Patrick's Day, Veteran's Day, Memorial Day, Flag Day, Women's Equality Day, and the Fiftieth Anniversary of D-Day). Of the 32 events that were linked to holidays, 16 occurred on July Fourth. As table 7 shows, events on July Fourth (generally children's parades) were less likely to be covered than other events, while events on other holidays were much more likely to be covered. We also created a dummy variable for annual events, that is, routinized or ritualized events with the same name, description, or organization held several years in a row at about the same time of year. These included such events as an Irish Dance on March 17, a Take Back the Night march, ceremonies around Martin Luther King Day, the gay pride march, the antiabortion movement's commemoration of *Roe v. Wade,* and a variety of annual fund-raising events. Consistent with the news-routine value of predictable events, the 49 annual events are much more likely to receive media coverage than other events.

Finally, the Wisconsin Blue Book was consulted to determine the exact dates on which the Wisconsin state legislature was in session in 1994, and a dummy variable indicates whether the legislature was in session on the particular date of the event. Table 7 indicates that whether the legislature is in session on a particular day appears to make little difference in the probability that a 1994 event is covered. However, the main legislative budget sessions are in odd-numbered years.

## Multivariate Analysis

Obviously, many of the factors predicting media coverage are correlated, and we need to use multivariate analysis to determine the net effects of each factor while controlling for the others. Two full models are shown in the appendix for each newspaper, giving the effects of all variables excluding and including the event typology. To make the models estimable for both newspapers, organizations were grouped as indicated in table 2, and vigils were combined with unpermitted protests. The reduced models shown in table 8 were constructed by selecting any variable that had a nontrivial coefficient (i.e., $P < .2$) in any of the four models after variables were eliminated by backward stepwise regression. The omitted categories for organizational sponsor, location, and event type were chosen to have moderate rates of news coverage, and various sensitivity analyses

TABLE 8

LOGIT REGRESSION OF NEWSPAPER COVERAGE

| | INCLUDING EVENT TYPES | | | | EXCLUDING EVENT TYPES | | | |
| | CT | | WSJ | | CT | | WSJ | |
| INDEPENDENT VARIABLE | Coefficient | SE | Coefficient | SE | Coefficient | SE | Coefficient | SE |
|---|---|---|---|---|---|---|---|---|
| Constant | -1.57*** | .58 | -1.80*** | .58 | -1.79+++ | .51 | -1 75**** | 51 |
| Size zero | -.24 | .58 | -.51 | .61 | -.81 | .52 | -1 00+ | 55 |
| Size small | -89 | .67 | -.85 | .65 | -1.61** | .62 | -1.17+ | 61 |
| Size large | 2.03*** | .50 | 1.97*** | .48 | 2.19*** | .48 | 2 20*** | 47 |
| Location | | | | | | | | |
| Downtown, outside[a] | -.61 | .48 | -1.15* | .47 | -.53 | .44 | -1.05* | 45 |
| Downtown, inside[b] | | | | | | | | |
| Inside at university | 1 06 | .81 | 1.40 | .78 | 84 | 80 | 1.23** | 77 |
| Library Mall area | -2.64** | .95 | -2 99** | .97 | -2 72** | .91 | -2.83+* | 92 |
| Not downtown | -1 79*** | .53 | -1 80** | .53 | -1.49** | .48 | -1.58*** | 49 |
| Organization type | | | | | | | | |
| Event specific | 1.18[1] | .69 | 1.26+ | .68 | 1.15+ | .65 | 1.27+ | 65 |
| Local SMO | -1.59 | 1.18 | -.59 | 1.08 | -.76 | 1.16 | -08 | 1.07 |
| National SMO | 2.76* | 1.21 | .53 | .89 | 3 09** | 1.18 | .72 | 91 |
| Local issue | .24 | .75 | .79 | .72 | .74 | .71 | 1.09 | 71 |
| National issue | .74 | 63 | 1.01 | .62 | 77 | .61 | 1.05+ | 61 |
| Recreational | 1 87** | 66 | 94 | .68 | 1.40* | .59 | .87 | 64 |
| Business | 2.12* | .83 | 1.67* | .79 | 2.17** | .78 | 1.85* | 75 |
| Nonprofit | 1.41 | .87 | 1 43+ | .86 | 1.46+ | .85 | 1.62+ | 84 |
| University students | 1.48+ | .85 | 2.03* | .82 | 1.72* | .82 | 2 08* | 81 |
| Business association | 3 62* | 1.40 | 2.34* | 1.10 | 3.21** | 1.22 | 2.53* | 99 |
| Public service | 1.11 | 97 | 1.79+ | .92 | .86 | .88 | 1.63+ | 88 |

Other:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Monday | 1.26* | .50 | 1.30** | .50 | 1.25** | .48 | 1.22* | .48 |
| Wednesday | −1.84* | .77 | −1.54* | .72 | −1.43* | .71 | −1.35+ | .70 |
| Saturday | −.57 | .39 | −.30 | .39 | −.73+ | .38 | −.35 | .37 |
| Message | −.10 | .51 | .06 | .50 | .85* | .40 | .64 | .40 |
| Conflict | 2.06*** | .56 | 1.85** | .55 | 1.59** | .51 | 1.38** | .51 |
| Vehicles | 1.10* | .56 | .82 | .54 | 1.11* | .54 | .63 | .53 |
| Nonlocal organizer | −1.46 | .98 | −1.64+ | .94 | −1.43 | .88 | −1.52+ | .88 |
| Amplifier | 1.12* | .56 | .89 | .53 | 1.10* | .48 | .95* | .48 |
| Annual | .50 | .53 | .76 | .50 | .54 | .50 | .76 | .48 |

Event type:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Rally | 2.92* | 1.18 | 1.84+ | 1.00 | | | | |
| Ceremony | .97 | .67 | .73 | .65 | | | | |
| Mixed | 1.26+ | .64 | 1.22+ | .63 | | | | |
| Commercial | .61 | .85 | 1.26 | .83 | | | | |
| Performance | −1.34+ | .68 | −.18 | .62 | | | | |

| | | | |
|---|---|---|---|
| Log likelihood | −135.23037 | −139.68721 | −143.8587 | −143.84157 |
| $\chi^2$ (df)[c] | 198.59 (32) | 167.94 (32) | 181.33 (27) | 159 63 |
| Probability[d] | .9998 | .9766 | .99995 | .9797 |

[a] Downtown outside = capitol steps, other downtown outside, or outside the university.

[b] Downtown inside = inside the capitol or inside other downtown buildings.

[c] Significance of $\chi^2$ test comparing models including and excluding event types is .1400 for *WSJ* and .0040 for *CT*.

[d] These entries show the probability of $\chi^2$ comparing the restricted model (this table) with the full model (app. table A1).

+ $P < .1$, two-tailed
* $P < .05$.
** $P < .01$.
*** $P < .001$.

71

were performed to be sure that the coefficients in the presented models represent major trends in the data and not outliers or small numbers of cases in a category. The selection of variables into the reduced models and the relative sizes of coefficients are consistent with the full models, and chi-square tests indicate that each reduced model fits as well as its full model.

*Newsworthiness.*—As expected, the size of an event is a major factor in the likelihood that it will receive coverage. Large events involving at least 500 people are much more likely to be covered by either newspaper than smaller events. However, size effects are less consistent below this threshold. The coefficient for very small events (1–15 people) is negative but is not generally significant, and the coverage events of size zero (i.e., displays and impersonal collections and distributions) is not significantly different than medium-size events.

But size is not all there is to newsworthiness. After size, the strongest effect on media coverage is the presence of conflict. This is not the same thing as disruption or disorder: neither the number of police, nor the subjective disorder code, nor being an unpermitted protest has any effect on news coverage, and the conflict effect is net of size. Net of conflict, there is some evidence that consensual message events receive somewhat more coverage than nonmessage events: the message coefficient is positive although not strong when event types are excluded.

More mundane newsworthiness factors also have effects. Events with amplifiers and events involving vehicles (i.e., street parades) receive more coverage, although these effects are stronger for the *Capital Times*. Since data on amplifiers come from permit records, the presence of amplified sound is probably an indication both of prior planning and of structuring the event in a way to speak to an audience. The effect of having a nonlocal organizer is consistently negative, which is consistent with local newspapers' assumed greater interest in events with local ties, although the effect is not strong enough to be significant.

*Sponsoring organization.*—Organizations clearly differed in their ability to attract newspaper attention to their events. Both newspapers give high rates of coverage to events sponsored by business associations and businesses and, to a weaker extent, the nonprofit institutions and organizations whose purpose is to put on an event. Although their gross rates of coverage are lower, University of Wisconsin student groups and recreational groups have positive effect coefficients because the events they sponsor are generally social or performance events with low coverage rates, and such events receive more coverage when they are sponsored by these groups than otherwise. Consistent with their respective editorial policies, the *Wisconsin State Journal* also is more likely to cover events

sponsored by consensual issue-oriented groups and the "public service" groups, while the *Capital Times* is especially likely to cover events sponsored by national SMOs.

*Location.*—It is a real estate truism that location is the key to property value. Location is important for receiving media coverage for an event as well. After multivariate controls, the original nine location codes could be grouped into the five shown in the table without loss of predictive value. The reference category in the table is the combination of inside the capitol and inside other downtown buildings, where both newspapers cover about one-third of the events. Net of controls for event characteristics, both newspapers covered events on the Library Mall or away from downtown much less often than those occurring in the capitol or other downtown buildings. The *Capital Times* had about the same rate of coverage for all downtown events, while the *Wisconsin State Journal* covered inside events at the university more and outside events downtown less than it covered events in the capitol or other downtown buildings. Location effects seem linked to news routines. Events in a small well-defined downtown area are easier to get to in a busy day than events in other parts of town. Furthermore, reporters are more likely already to be downtown for other reasons. While collecting data for this project, we observed an unpermitted blockade of the legislative chambers over an antiabortion bill. The first reporters on the scene were those who were already in the building for another purpose, presumably because their beat is the capitol. Other reporters and television crews (who were called by the protesters) arrived later.

The importance of location can be most strongly seen in the contrast between ceremonies and "mixed" events. Neither type of event is typically thought of as protest, and both generally carry consensual rather than conflictual content. The "mixed" events should generally be more newsworthy than ceremonies, as they are much larger than ceremonies and involve some sort of interesting activity. But 59% of the ceremonies received coverage versus only 40% of the mixed events. It turns out that this difference is entirely accounted for by the location of the event. Ceremonies are more likely to occur in the capitol, while mixed events are more likely to occur away from downtown or on the Library Mall. In fact, at any given location, a mixed event is *more* likely to receive coverage than a ceremony.

*Timing and routine.*—The bivariate effect for annual events is largely explained when other variables are controlled, although it remains positive, especially for the *Wisconsin State Journal.* Although the proportions of events that are annual differ across event types, coverage for annual events is consistently higher across the types (except for ceremonies, where

there is little difference). When reporters can plan ahead for the event and rely on templates and understandings from prior years, any kind of event is more likely to be covered, but this is a relatively weak effect.

It is less clear why events on Monday are much more likely to be covered, and events on Wednesday and Saturday are somewhat less likely to be covered, but this different effect is consistent between the two newspapers and remains after multivariate controls. This is most likely an effect of the news hole (the amount of space available on a given day for news), as the newspapers' news holes do differ by day of the week because of regular fluctuations in advertising. However, there is no simple correspondence between the day an event occurs and the day it is reported.

*Event forms.*—Comparisons of fit statistics for model 1 and model 2 indicate that event forms affect media coverage over and above the other characteristics of events, and that this effect is strong and significant for the *Capital Times* in both the full and reduced models. The main factor is that the *Capital Times* was especially likely to cover rallies, although the *Wisconsin State Journal* also has a marginally significant tendency to cover rallies as well. Additionally, mixed events and, to a lesser extent, ceremonies were covered more than their other characteristics (principally location) would predict. Conversely, performances were less likely to be covered, especially by the *Capital Times*. The relative coverage rates of particular protest forms needs to be viewed as conditional. Prior research (McCarthy, McPhail, and Smith 1996) found that, in Washington, D.C., marches were more likely to be covered than rallies; our sample of marches and rallies in Madison is relatively small. However, everyone seems to find that vigils receive very little news coverage.

DISCUSSION

Research in this tradition began by asking which protest events receive news coverage. This question is important in two ways, both because most research on protest events relies on news reportage of events as data and because protest events are an alternate path to influence for people who feel aggrieved by social policies. But an adequate answer to the question of *which* protest events receive coverage requires answering a prior question: How do protest events fare in competition with other events for space in the news hole? Is protest an effective way of bringing issues to the public sphere by way of news coverage? We can answer this question only by comparing protest events to other kinds of events. When we make the comparison, it is clear that protest is quite effective in competition with other similar events for news space. Local newspapers may seem to be full of small stories about social events, performances, and athletic competitions, but police records make it clear that there are many more of these

events than other kinds of events and that events purely for entertainment, with no message, are much less likely to be covered in newspapers than events with messages.

The results of our research allow us to place public events on a coverage scale. On the one end are promotional events for local businesses and business sectors, which have extremely high rates of news coverage, often in the "business" section which is explicitly oriented to business interests. On the other end of the scale are the purely social and entertainment events which are ends in themselves, have no particular relevance to people who are not present at the event, and have little or no "news value" by any journalistic standard. These events are quite unlikely to be covered, unless they are large. Athletic events and most message events fall between these extremes. Regardless of editorial position, this scale and the location of most event types on it is similar for the two newspapers.

With this scale established, the coverage of events carrying protest content can be calibrated: rallies are near the high end, while marches, speeches, and mixed events are in the middle (receiving coverage comparable to athletic events and parades). Unpermitted protests have relatively low rates of coverage (comparable to the rates for displays and collections), although they receive more news attention than social and entertainment events. The more liberal newspaper gave considerably greater coverage to rallies than the more conservative newspaper, but even the more conservative newspaper gave rallies and other protest events relatively high coverage as compared with other event types. In fact, the data suggest that there is little bias against protest. Instead, it is social events and small performances that are especially likely to be omitted from the news record.

But the omission of events that are not, after all, trying to communicate a message to a larger public seems less noteworthy than the omission of nonconflictual message events. Instead, if there is a bias, it appears to be against messages that lack conflict. If conflict exists about an event or an issue connected to it, the event is much more likely to receive news coverage. The data suggest that it is drama, not novelty, that makes an event "news." In this respect, it is worth noting that athletic events, which involve the drama of competition, can be "news" in a way that performances or social events cannot. If conflict or drama is news, ceteris paribus, social movements have an advantage over other kinds of groups in obtaining news coverage. The contrast between conflictual and consensual issues is enormous. Fourteen (88%) of the 16 rallies, marches, vigils, ceremonies, and speeches around a conflictual issue received media coverage, while only 46% of the 54 rallies, marches, vigils, ceremonies, and speeches involving consensual issues received media coverage. Displays are much less likely to be covered than events, but 42% of the seven displays involving conflict received some coverage versus only 18% of the 22 involving

a consensual issue. This is not affected by editorial policy. The *Wisconsin State Journal* shows the same pattern of considerably higher coverage for conflictual than consensual messages, even though its coverage of both kinds of message events is lower than the *Capital Times*.

The type of organizational sponsor also has a major effect on news coverage. Business associations and businesses were very successful in receiving coverage regardless of the editorial stance of the newspaper and net of other event characteristics. Net of other event characteristics, non-profit institutions, event-specific groups, service organizations, university students, and recreational groups also were treated favorably by both newspapers. It is especially worth calling attention to the comparison be-tween the social movement organizations and the national and local issue-oriented groups. The issue-oriented groups are a lot like movement organi-zations, in that they are voluntary associations that seek to educate or persuade the public to new ideas and to make claims that, if realized, would affect people other than themselves. In fact, issue-oriented groups often advocate for increased public funding or other legislation relevant to their issue. Where they differ from movement organizations is in their nonconflictual and relatively nonpolitical appeal to public issues or con-cerns. They frame their issues as charitable or educational or consensual public concerns. They frame themselves and are framed by others (includ-ing movement activists and social movement scholars) as very different from social movements, as insiders seeking to do good rather than outsid-ers seeking controversial change. Both newspapers covered around a third of the events sponsored by issue-oriented groups, which is about the same coverage they gave to local movement organizations. When multivariate controls are introduced, the effect coefficient becomes negative for local SMOs and positive for the local and national issue groups. This means that, net of the strong positive effect of conflict, local SMOs receive less coverage than other organizations, and issue groups receive more. How-ever, these effects are generally weak and nonsignificant. Net of controls, the more moderate *Wisconsin State Journal* favored the issue groups—especially national issue groups-more than the *Capital Times* did, but these effects are not large enough to be significant. The big difference editorial position made was in the liberal *Capital Times*'s higher coverage of national social movement organizations. These coverage rates suggest that social movement organizations can be "insiders," at least for some media organizations.

Location and the spatial dimension of action have also been insuffi-ciently appreciated in earlier discussions. Just as early research has dem-onstrated the regional bias of any news outlet, this research has demon-strated the existence of smaller-scale geographic biases in news coverage within one small city. Events occurring away from the downtown area

had much less chance of receiving news coverage than comparable events occurring near the state capitol. "Public" events are intended to be observed by strangers and, ideally, publicized through the mass media. Most public events in Madison tend to occur in well-defined and narrowly limited public spaces, and news coverage focused on those limited areas. Events occurring outside these well-defined areas received much less coverage. Additionally, the data suggest that there is some avoidance of the Library Mall area by news reporters, which is viewed as a "student" domain. Several news people mentioned the lack of news value of "another protest by the same group of people on Library Mall." Urban geographers and anthropologists have been calling attention to the spatial organization of urban life, to the ways in which different groups of people frequent different areas of town and use public space differently. When people differ in their access to the places frequented or considered significant by news reporters, they will differ in their ability to reach a larger audience through public events. Additionally, mundane factors that affect newsworthiness or news routines appear to make some difference in coverage, including the day of the week, the presence of vehicles, the use of an amplifier, and the absence of a local organizer.

The lessons of these results for activists seeking media coverage seem fairly clear. Events involving conflictual messages fare rather well in local news coverage, particularly if they occur in the right places and have an organizational sponsor with positive ties to the news media. Activists seeking to influence the public or to mobilize action on a consensual basis, on the other hand, appear to be fighting an uphill battle. However, contrary to common assumptions, there appears to be no premium for disruption or surprise in obtaining news coverage: unpermitted protests were not especially likely to be covered, and neither the number of police nor the subjective measure of disorder predicted news coverage. Timing the event to meet news deadlines seems less important for newspaper coverage than recognizing daily differences in news hole sizes.

The lessons for protest event researchers are also fairly clear. Protest does appear to be "newsworthy" and to have a reasonably high probability of coverage. However, there is a distinct bias in the news record not only toward larger events (which has been well documented) but toward more established groups who present their concerns in predictable ways in predictable places located downtown, where the reporters are. The genuinely marginal people, who put on their events in their own spaces, are much more likely to be missing from the news record.

Finally, these results have important implications for our understanding of the role of media coverage of public events in creating the public sphere of democratic discourse. In this discussion, it is important to remember the scope of the research, that we have no way to assess the

coverage of public events against the routine access of political or economic elites to news reporters. Additionally, the police data do not record most events in private spaces. With these caveats in mind, there are three main conclusions to be drawn about the impact of public events on the public sphere. First, although there has been considerable discussion of the timing of events relative to news deadlines, these data suggest that much more attention should be given to *where* events occur and to the spatial as well as social accessibility of events to news reporters. In particular, the locations that received media coverage (the capitol area and inside university buildings) are not the locations where a majority of the population goes (i.e., shopping malls and sports arenas). This implies that public events can be oriented either to those physically present at the event, or to the mass media, but that it is relatively difficult to do both.

Second, the data suggest that news organizations are not necessarily homogeneous, particularly if they are in a competitive market. Both newspapers, regardless of editorial position, covered events of importance to the business community. But this does not mean that their editorial positions are irrelevant to their decisions about which kinds of other events are important and worthy of coverage. The more liberal newspaper gave substantially more coverage to rallies and to events sponsored by national SMOs than did the more moderate newspaper. Although some journalistic norms of "objective" news reporting imply that editorial position should be uncorrelated with news coverage, traditional conceptions of the role of a free press in democratic discourse in the public sphere suggest the value of diverse and competing definitions of which issues are important to the public. If staging public events is a mechanism for bringing issues into the public sphere that are not already being discussed, it follows that there cannot be a priori objective standards for which of these issues are important enough to receive media attention.

Finally, the results suggest that the drama of controversy and polarization is central to what is understood by reporters and the general public as "interesting" or "newsworthy," and that the way an issue is framed is more salient than the novelty or disruptiveness of an event, especially if the disruption remains relatively small. Although they have not had comparable data on the coverage of large numbers of events, this news preference for conflict over cooperation is well recognized among media scholars (Shah and Thornton 1994; Shoemaker and Resse 1991; Thornton and Shah 1996). At least by way of a public event, it appears relatively difficult to obtain media coverage to promote good health, education, charity, volunteering, or positive community relationships. Although public events are not the only means by which issues enter the public sphere, the results in this article strongly support arguments that the information the public receives from the media about social issues is slanted toward

conflict and controversy. It may be useful in closing to consider this contentious "public sphere" one encounters in the newspaper with the full range of events actually occurring in the "public square," where people are in physical proximity to one another. In the public square, people not only contend, they also socialize, entertain each other, raise money for charity, and seek to influence each other in nonconflictual ways. This image of positive sociability is at variance with many images of public life, and it is possible that Madison is unusually civil. But the data suggest that it is also possible that the images of public interactions have been too heavily distorted by reliance on media coverage as a source of information about those interactions.

APPENDIX

Full Models

TABLE A1

LOGIT REGRESSION OF NEWSPAPER COVERAGE

| | EXCLUDING EVENT TYPES | | | | INCLUDING EVENT TYPES | | | |
| | CT | | WSJ | | CT | | WSJ | |
| INDEPENDENT VARIABLE | Coefficient | SE | Coefficient | SE | Coefficient | SE | Coefficient | SE |
|---|---|---|---|---|---|---|---|---|
| Constant | −2.78* | 1.27 | −3.29* | 1.31 | −4.15* | 1.98 | −2.99 | 1.99 |
| Size zero | −.99 | .65 | −1.00 | .70 | −.21 | 1.65 | .08 | 1.86 |
| Size small | −1.38* | .68 | −1.18+ | .68 | −.62 | .75 | −.71 | .75 |
| Size large | 2.22*** | .51 | 2.31*** | .50 | 2.08*** | .57 | 1.82** | .53 |
| Organization type: | | | | | | | | |
| None | −.78 | .81 | −.90 | .82 | −.87 | .96 | −.91 | .94 |
| Event specific | .99 | .84 | 1.17 | .88 | .70 | .98 | 1.00 | 1.00 |
| Local SMO | −1.05 | 1.31 | −.59 | 1.24 | −2.08 | 1.60 | −1.62 | 1.47 |
| National SMO | 2.12* | 1.01 | .99 | 1.00 | 1.54 | 1.05 | .88 | 1.03 |
| Local issue | .49 | .82 | .99 | .82 | −.26 | .91 | .63 | .89 |
| National issue | .73 | .79 | 1.09 | .81 | .45 | .88 | .85 | .90 |
| Religious | −.53 | 1.06 | .10 | 1.04 | −.92 | 1.19 | −.45 | 1.17 |
| Occupational | .72 | 1.23 | −1.71 | 1.64 | .33 | 1.34 | −2.25 | 1.86 |
| Recreational | 1.20 | .83 | .87 | .87 | 1.52 | .97 | .54 | .99 |
| Business (particular) | 1.87+ | .96 | 1.76+ | .96 | 1.71+ | 1.04 | 1.47 | 1.03 |
| Youth group, school | −.53 | 1.32 | −.52 | 1.33 | −.23 | 1.53 | −.74 | 1.44 |
| University | −.08 | 1.19 | −.21 | 1.26 | .19 | 1.43 | .63 | 1.53 |
| Nonprofit | 1.40 | .99 | 1.50 | 1.00 | 1.25 | 1.05 | 1.30 | 1.05 |
| University student group | .73 | 1.16 | 1.11 | 1.14 | .80 | 1.38 | 1.18 | 1.31 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Business association | 2.88* | 1 36 | 2.43* | 1.19 | 3.52* | 1.55 | 2.17 | 1.38 |
| Neighborhood | .57 | 1 32 | .23 | 1.38 | 1 60 | 1.73 | 1.22 | 1.73 |
| Public service | 74 | 1.05 | 1.72 | 1.04 | 99 | 1.17 | 1.79 | 1.13 |
| **Location:** | | | | | | | | |
| Inside capitol | .58 | .59 | .80 | 59 | 1.03 | .65 | 1.19+ | .64 |
| Inside other downtown | .73 | .96 | 1.02 | .99 | .73 | 1 06 | .78 | 1 11 |
| Inside at university | 1.61 | 1.10 | 2.64* | 1.11 | 1.95 | 1.20 | 3.06* | 1.18 |
| Inside, not downtown | -.18 | .99 | -.38 | 1.04 | -.16 | 1.06 | -.59 | 1.15 |
| Outside other downtown | .26 | .67 | -.53 | 69 | .74 | .76 | -.44 | .79 |
| Library Mall | -2 07* | 1.03 | -1.72+ | 1.03 | -1.50 | 1 17 | -1.50 | 1.18 |
| Outside at university | 89 | 1.02 | .43 | 1.05 | 1.72 | 1 29 | .71 | 1.23 |
| Outside, not downtown | -1.18+ | .66 | -.90 | .66 | -.95 | .77 | -.91 | .76 |
| **Time of day:** | | | | | | | | |
| All day | -.22 | .56 | 18 | 57 | -.35 | .60 | .10 | .62 |
| Morning | -.27 | .57 | -.09 | .60 | -.19 | .61 | -.09 | .65 |
| Midday | -.04 | .59 | .39 | .59 | -.10 | .66 | 37 | 65 |
| Afternoon | 17 | 60 | .43 | .61 | .05 | .67 | .48 | .67 |
| **Day of week:** | | | | | | | | |
| Sunday | -02 | 79 | -.32 | 82 | -.11 | 84 | -.40 | .87 |
| Monday | .99 | .80 | 1.10 | 82 | 79 | .84 | 1.10 | 87 |
| Tuesday | 05 | .82 | .34 | .86 | -04 | .89 | .42 | .93 |
| Wednesday | -1 33 | .90 | -1.44 | .89 | -2.01* | 1.01 | -1 58 | 98 |
| Friday | -02 | .77 | -.14 | .81 | -.54 | .82 | -.35 | .86 |
| Saturday | -71 | 72 | -.51 | .75 | -.66 | .76 | -.43 | 82 |
| **Other:** | | | | | | | | |
| July Fourth | -.79 | 1.30 | -.73 | 1.38 | -1.01 | 1.50 | -.61 | 1 54 |
| Other holiday | .95 | 79 | .33 | .78 | .95 | .86 | .86 | .85 |
| Message | 98* | 48 | .61 | 48 | -.46 | .66 | -.10 | .67 |
| Conflict | 1.62* | .64 | 1.52* | .64 | 2.41** | 78 | 2.34*** | .80 |
| Disorder | -.11 | .89 | .45 | .88 | .23 | 1.05 | 1.31 | 1.20 |
| Vehicles | 1.22* | .61 | .65 | .60 | 1.66* | .73 | .74 | .69 |

TABLE A1 (Continued)

| INDEPENDENT VARIABLE | EXCLUDING EVENT TYPES | | | | INCLUDING EVENT TYPES | | | |
|---|---|---|---|---|---|---|---|---|
| | CT | | WSJ | | CT | | WSJ | |
| | Coefficient | SE | Coefficient | SE | Coefficient | SE | Coefficient | SE |
| Other. | | | | | | | | |
| Nonlocal organizer | −1.18 | .97 | −1.39 | 1 02 | −1 29 | 1.09 | −1.47 | 1.11 |
| Amplifier | 1.05+ | 56 | 87 | .55 | .98 | 63 | 72 | 60 |
| Annual | 30 | 56 | 84 | 54 | .58 | 63 | .93 | .59 |
| Legislature in session | −.24 | 50 | −.16 | 51 | −.30 | 55 | −.10 | .56 |
| N of police | 12 | 18 | 22 | 18 | 12 | 19 | 28 | 19 |
| N events in 17 days | 21 | 18 | 23 | 18 | 52 | 35 | 25 | 34 |
| Event type: | | | | | | | | |
| Rally | | | | | 4.69** | 1 67 | 1.52 | 1 44 |
| March | | | | | 1.76 | 1 56 | −.03 | 1.49 |
| Ceremony | | | | | 2 11+ | 1 17 | −.03 | 1.19 |
| Speech | | | | | 1.30 | 1.40 | −1 25 | 1 64 |
| Protest or vigil | | | | | .79 | 1.59 | −1.78 | 1 71 |
| Display | | | | | .72 | 2.00 | −1.48 | 2 29 |
| Collection, distribution | | | | | 1.82 | 1 88 | −.12 | 2.01 |
| Mixed | | | | | 1.93+ | .99 | .98 | 97 |
| Commercial | | | | | .81 | 1.05 | 1.12 | 1.10 |
| Social | | | | | −1 58 | 1 67 | −1.40 | 1.60 |
| Performance | | | | | −.39 | .95 | −.50 | 97 |
| Parade | | | | | −1 13 | 1.40 | −1 06 | 1.31 |
| Log likelihood | −141.21604 | | −138.17967 | | −130.23979 | | −131.36407 | |
| χ² (df) | 186.62 (50) | | 170.95 (50) | | 208 57 (62) | | 184.58 (62) | |
| Probability^a | | | | | .0381 | | 3250 | |

NOTE —Omitted dummy variables are medium size, government organization type, capitol steps location, evening, Thursday, and athletic event  CT = Capital Times, WSJ = Wisconsin State Journal
^a This test compares models including event types with models excluding event types
+ P < 1, two-tailed, * P < 05, ** P < .01, *** P < 001

REFERENCES

Alario, Margarita. 1994. "Environmental Destruction and the Public Sphere: On Habermas's Discursive Model and Political Ecology." *Social Theory and Practice* 20:327–41.

Altheide, David L., and Paul K. Rasmussen. 1976. "Becoming News: A Study of Two Newsrooms." *Sociology of Work and Occupations* 3:223–46.

Bagdikian, Benjamin. 1983. *The Media Monopoly*. Boston: Beacon Press.

Beharrell, P. , and G. Philo, eds. 1977. *Trade Unions and the Media*. London: Macmillan.

Bennett, W. Lance. 1980. "Myth, Ritual, and Political Control." *Journal of Communication* 30:166–79.

Bowers, D. 1967. "A Report on Activity by Publishers in Directing Newsroom Decisions." *Journalism Quarterly* 1:43–52.

Brown, Elsa Barkley, and Gregg D. Kimball. 1995. "Mapping the Terrain of Black Richmond." *Journal of Urban History* 21:296–346.

Chaffee, Lyman. 1993. "Dramaturgical Politics: The Culture and Ritual of Demonstrations in Argentina." *Media Culture and Society* 15:113–35.

Cohn, Steven F., and James E. Gallagher. 1984. "Gay Movements and Legal Change: Some Aspects of the Dynamics of a Social Problem." *Social Problems* 32:72–86.

Costain, Anne N. 1992. *Inviting Women's Rebellion: A Political Process Interpretation of the Women's Movement*. Baltimore: Johns Hopkins University Press.

Danzger, M. Herbert. 1975. "Validating Conflict Data." *American Sociological Review* 40:570–84.

Della Porta, Donatella. 1996a. "Sapere di polizia e ordine pubblico: Alcune riflessioni sul caso italiano (Police knowledge and public order: Some reflections on the Italian case)." *Polis* 10:361–82.

———. 1996b. "Social Movements and the State: Thoughts on the Policing of Protest." Pp. 62–92 in *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings,* edited by Doug McAdam, John D. McCarthy, and Mayer N. Zald. New York: Cambridge University Press.

Della Porta, Donatella, and Herbert Reiter, eds. 1998. *Policing Protest: The Control of Mass Demonstrations in Western Democracies*. Minneapolis: University of Minnesota Press.

Downing, J. 1980. *The Media Machine*. London: Pluto Press.

Downs, Anthony. 1972. "Up and Down with Ecology—the Issue Attention Cycle." *Public Interest* 28:38–50.

Drucker, Susan J., and Gary Gumpert. 1996. "The Regulation of Public Social Life: Communication Law Revisited." *Communication Quarterly* 44:280–96.

Eisinger, Peter K. 1973. "The Conditions of Protest Behavior in American Cities." *American Political Science Review* 67:11–28.

Fillieule, Olivier. 1998. " 'Plus ça change, moins ça change': Demonstrations in France during the Nineteen-Eighties." Pp. 199–226 in *Acts of Dissent: New Developments in the Study of Protest,* edited by Dieter Rucht, Ruud Koopmans, and Friedhelm Neidhardt. Berlin: Edition Sigmas Rainer Bohn Verlag.

Franzosi, Roberto. 1987. "The Press as a Source of Socio-Historical Data: Issues in the Methodology of Data Collection from Newspapers." *Historical Methods* 20:5–15.

Gamson, William A. 1995. "Constructing Social Protest." Pp. 85–106 in *Social Movements and Culture,* edited by Hank Johnston and Bert Klandermans. Minneapolis: University of Minnesota Press.

Gamson, William A., David Croteau, William Hoynes, and Theodore Sasson. 1992. "Media Images and the Social Construction of Reality." *Annual Review of Sociology* 18:373–93.

Gamson, William A., and David S. Meyer. 1996. "Framing Political Opportunity." Pp 273–90 in *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings*, edited by Doug McAdam, John D. McCarthy, and Mayer N Zald. New York: Cambridge University Press

Gamson, William A., and Andre Modigliani. 1989. "Media Discourse and Public Opinion on Nuclear Power: A Constructionist Approach." *American Journal of Sociology* 95:1–37.

Gamson, William A., and Gadi Wolfsfeld. 1993. "Movements and Media as Interacting Systems." *Annals of the American Academy of Political and Social Science* 528:114–25.

Gans, Herbert. 1980. *Deciding What's News: A Study of CBS Evening News, NBC Nightly News, Newsweek and Time*. New York: Vintage.

Gitlin, Todd. 1980. *The Whole World Is Watching: Mass Media in the Making and Unmaking of the New Left*. Berkeley and Los Angeles· University of California Press

Goren, Dina. 1980. "The News and Foreign Policy: An Examination of the Impact of the News Media on the Making of Foreign Policy." *Research in Social Movements, Conflicts and Change* 3:119–41.

Gurr, T. R. 1968. "A Causal Model of Civil Strife: A Comparative Analysis Using New Indices." *American Political Science Review* 62:1104–24.

Hartmann, P. 1979. "News and Public Perceptions of Industrial Relations." *Media, Culture, and Society* 1:255–70.

Hellman, Judith Adler. 1987. *Journeys among Women: Feminism in Five Italian Cities*. New York: Oxford University Press.

Herman, Edward, and Noam Chomsky. 1988. *Manufacturing Consent*. New York: Pantheon.

Hilgartner, Stephen, and Charles L. Bosk. 1988. "The Rise and Fall of Social Problems: A Public Arenas Model." *American Journal of Sociology* 94:53–78.

Hocke, Peter 1998. "Determining the Selection Bias in Local and National Newspaper Reports on Protest Events." Pp. 131–63 in *Acts of Dissent*, edited by Dieter Rucht, Ruud Koopmans, and Friedhelm Neidhardt. Berlin: Edition Sigma Rainer Bohn Verlag.

Jacobs, Ronald N. 1996. "Producing the News, Producing the Crisis: Narrativity, Television and News Work." *Media Culture and Society* 18:373–97.

Jenkins, J Craig, and Craig M. Eckert. 1986. "Channeling Black Insurgency: Elite Patronage and Professional Social Movement Organizations in the Development of the Black Movement." *American Sociological Review* 51:812–29.

Jenkins, J. Craig, and Charles Perrow. 1977. "Insurgency of the Powerless: Farm Worker Movements (1946–1972)." *American Sociological Review* 42·249–68.

Kielbowicz, Richard B., and Clifford Scherer. 1986. "The Role of the Press in the Dynamics of Social Movements." *Research in Social Movements, Conflicts and Change* 9:71–96.

Koenen, Elmar J 1996. "Die Wiederkehr der Offentlichkeit (The Return of the Public Sphere)." *Soziologische Revue* 19:36–41.

Koopmans, Ruud. 1993. "The Dynamics of Protest Waves· West Germany, 1965 to 1989." *American Sociological Review* 58:637–58.

Kriesi, Hanspeter, Ruud Koopmans, Jan Willem Duyvendak, and Marco G. Giugni 1995. *New Social Movements in Western Europe: A Comparative Analysis*. Minneapolis: University of Minnesota Press.

Lee, M. A., and N. Solomon. 1990. *Unreliable Sources: A Guide to Detecting Bias in News Media.* New York: Carol Publishing Group.

Lees, Lynn Hollen. 1994. "Urban Public Space and Imagined Communities in the 1980s and 1990s." *Journal of Urban History* 20:443–65.

Lester, Marilyn. 1980. "Generating Newsworthiness The Interpretive Construction of Public Events." *American Sociological Review* 45:984–94.

Lieberson, Stanley, and Arnold R Silverman. 1965. "The Precipitants and Underlying Conditions of Race Riots." *American Sociological Review* 30:887–98.

Lipsky, Michael. 1968. "Protest as a Political Resource." *American Political Science Review* 62:1144–58

Lofland, John, and Michael Fink. 1982 *Symbolic Sit-Ins: Protest at the California Capitol.* Washington, D.C.: University Press of America

Maney, Gregory M., and Pamela E. Oliver 1998. "Finding Event Records: Timing, Searching, Sources " Working paper University of Wisconsin, Madison.

McAdam, Doug. 1982. *Political Process and the Development of Black Insurgency, 1930–1970.* Chicago: University of Chicago Press.

———. 1983. "Tactical Innovation and the Pace of Insurgency." *American Sociological Review* 48 735–54.

McCarthy, John D., David W. Britt, and Mark Wolfson. 1991. "The Institutional Channeling of Social Movements by the State in the United States." *Research in Social Movements, Conflicts and Change* 13:45–76

McCarthy, John D., and Clark McPhail. 1997. "The Institutionalization of Protest " Pp 83–110 in *A Movement Society? Contentious Politics for a New Century,* edited by David Meyer and Sidney Tarrow Boulder, Colo. Rowland & Littlefield.

McCarthy, John D , Clark McPhail, and John Crist. 1998. "The Emergence and Diffusion of Public Order Management Systems· Protest Cycles and Police Response." In *Globalization and Social Movements,* edited by Hanspeter Kriesi, Donatella della Porta, and Dieter Rucht. Houndsmill: Macmillan.

McCarthy, John D., Clark McPhail, and Jackie Smith. 1996 "Images of Protest: Dimensions of Selection Bias in Media Coverage of Washington Demonstrations, 1982 and 1991 " *American Sociological Review* 61:478–99.

McCarthy, John D., Clark McPhail, Jackie Smith, and Louis J. Crishock 1998. "Electronic and Print Media Representations of Washington, D.C. Demonstrations, 1982 and 1991: A Demography of Description Bias." Pp. 113–30 in *Acts of Dissent: New Developments in the Study of Protest,* edited by Dieter Rucht, Ruud Koopmans, and Friedhelm Neidhardt. Berlin Edition Sigmas Rainer Bohn Verlag

McCarthy, John D , and Mayer N. Zald. 1977. "Resource Mobilization and Social Movements: A Partial Theory." *American Journal of Sociology* 82:1212–41.

McPhail, Clark, David Schweingruber, and John D. McCarthy. 1998. "Policing Protest in the United States, 1960–1995." Pp. 49–69 in *Policing Protest: The Control of Mass Demonstrations in Western Democracies,* edited by Donatella della Porta and Herbert Reiter. Minneapolis: University of Minnesota Press

McPhail, Clark, and Ronald T. Wohlstein. 1986. "Collective Locomotion as Collective Behavior." *American Sociological Review* 51.447–63.

Meyer, David S. 1993. "Institutionalizing Dissent: The United States Structure of Political Opportunity and the End of the Nuclear Freeze Movement." *Sociological Forum* 8:157–79.

Meyer, David S., and Joshua P Gamson. 1993. "The Challenge of Cultural Elites: Celebrities and Social Movements." *Sociological Inquiry* 65:181–206.

Molotch, Harvey. 1979 "Media and Movements." Pp. 71–93 in *The Dynamics of Social Movements,* edited by Mayer N. Zald and John D McCarthy. Cambridge, Mass.: Winthrop

Molotch, Harvey, and Marilyn Lester. 1974. "News as Purposive Behavior. On the

Strategic Use of Routine Events, Accidents, and Scandals." *American Sociological Review* 39:101–12.

Morley, David. 1976. "Industrial Conflict and the Mass Media." *Sociological Review* 24:245–68.

Mueller, Carol. 1997a. "International Press Coverage of East German Protest Events, 1989." *American Sociological Review* 62:820–32.

———. 1997b. "Media Measurement Models of Protest Event Data." *Mobilization* 2: 165–84.

Olien, Clarice N., Phillip J. Tichenor, and George A. Donahue. 1992. "Media Coverage and Social Movements." In *Information Campaigns: Balancing Social Values and Social Change,* edited by C. T. Salmon. Newbury Park, Calif.: Sage.

Oliver, Pamela E. 1989 "Bringing the Crowd Back In: The Nonorganizational Elements of Social Movements " *Research in Social Movements, Conflicts and Change* 11:1–30.

Oliver, Pamela E. , and Gregory M. Maney. 1998. "Making the Local News: Police and Media Records of Protest Events in a Small City." Paper presented at the annual meeting of the American Sociological Association, San Francisco.

Oliver, Pamela E., and Gerald Marwell. 1992. "Mobilizing Technologies for Collective Action." Pp. 251–72 in *Frontiers in Social Movement Theory,* edited by Aldon Morris and Carol McClurg Mueller. New Haven, Conn.: Yale University Press.

Olzak, Susan. 1992. *The Dynamics of Ethnic Competition and Conflict.* Stanford, Calif.: Stanford University Press.

Olzak, Susan, Suzanne Shanahan, and Elizabeth West. 1994. "School Desegregation, Interracial Exposure, and Antibusing Activity in Contemporary Urban America." *American Journal of Sociology* 100:196–241.

Parenti, Michael. 1986. *The Politics of the Mass Media.* New York: St. Martin's Press.

Ray, Raka. 1993. "Public Agendas and Women's Interests: Organizing Women in Two Indian Cities." Ph.D. dissertation. University of Wisconsin, Department of Sociology.

Rothman, Franklin Daniel. 1993. "Political Process and Peasant Opposition to Large Hydroelectric Dams: The Case of the Rio Uruguai Movement in Southern Brazil, 1979 to 1992." Ph.D. dissertation. University of Wisconsin, Department of Sociology.

Rucht, Dieter, and Friedhelm Neidhardt. 1998. "Methodological Issues in Collecting Protest Event Data: Units of Analysis, Sources and Sampling, Coding Problems." Pp. 65–89 in *Acts of Dissent: New Developments in the Study of Protest,* edited by Dieter Rucht, Ruud Koopmans, and Friedhelm Neidhardt. Berlin: Edition Sigmas Rainer Bohn Verlag.

Ryan, Charlotte. 1991. *Prime Time Activism.* Boston: South End Press.

Salzman, Jason. 1998. *Making the News: A Guide for Nonprofits and Activists.* Boulder, Colo.: Westview Press.

Schulz, Winfried. 1982. "News Structure and People's Awareness of Political Events." *Gazette* 30:139–53.

Sebastiani, Chiara. 1997. "Spazio e sfera pubblica: La politica nella citta (Space and the public sphere: Politics within the city)." *Rassegna Italiana di Sociologia* 38:223–43.

Shah, Hemant, and Michael C. Thornton. 1994. "Racial Ideology in U.S. Mainstream News Magazine Coverage of Black-Latino Interaction, 1980–1992." *Critical Studies in Mass Communication* 11:141–61.

Shoemaker, Pamela J., and Stephen D. Resse. 1991. *Mediating the Message: Theories of Influences on Mass Media Content.* New York: Longman.

Shorter, Edward, and Charles Tilly. 1974. *Strikes in France.* Cambridge: Cambridge University Press.

Sigelman, Lee. 1973 "Reporting the News: An Organizational Analysis." *American Journal of Sociology* 79:132–51.

Snyder, David, and William R. Kelly. 1977. "Conflict Intensity, Media Sensitivity and the Validity of Newspaper Data." *American Sociological Review* 42:105–23.

Spilerman, Seymour. 1970. "The Causes of Racial Disturbances: A Comparison of Alternative Explanations." *American Sociological Review* 35:627–49.

———. 1976. "Structural Characteristics of Cities and the Severity of Racial Disorders." *American Sociological Review* 41:771–93

Staggenborg, Suzanne. 1988. "The Consequences of Professionalization and Formalization in the Pro-Choice Movement." *American Sociological Review* 53:585–605.

Tarrow, Sidney. 1994. *Power in Movement: Social Movements, Collective Action and Mass Politics in the Modern State.* Cambridge: Cambridge University Press.

———. 1998. "Studying Contentious Politics: From Event-ful History to Cycles of Collective Action." Pp. 33–64 in *Acts of Dissent: New Developments in the Study of Protest,* edited by Dieter Rucht, Ruud Koopmans, and Friedhelm Neidhardt. Berlin: Edition Sigmas Rainer Bohn Verlag.

Thornton, Michael C., and Hemant Shah. 1996. "U.S. News Magazine Images of Black–Asian American Relationships, 1980–1992." *Communication Review* 1:497–519.

Tilly, Charles. 1978. *From Mobilization to Revolution.* Reading, Mass.: Addison-Wesley

———. 1986. *The Contentious French: Four Centuries of Popular Struggle.* Cambridge, Mass.: Belknap Press.

Zald, Mayer N. 1996. "Culture, Ideology, and Strategic Framing." Pp. 261–74 in *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings,* edited by Doug McAdam, John D. McCarthy, and Mayer N. Zald  New York: Cambridge University Press.

# Is Social Capital Declining in the United States? A Multiple Indicator Assessment[1]

Pamela Paxton
*Ohio State University*

> Despite a great deal of interest in a possible decline of social capital
> in the United States, scholars have not reached a consensus on the
> trend. This article improves upon previous research by providing a
> model of social capital that has explicit links to theories of social
> capital and that analyzes multiple indicators of social capital over
> a 20-year period. The results do not consistently support Putnam's
> claim of a decline in social capital, showing instead some decline in
> a general measure of social capital, a decline in trust in individuals,
> no general decline in trust in institutions, and no decline in associa-
> tions.

Concern about a decline in community is a recurring theme in classical
and contemporary sociology. In fact, it could be argued that the birth of
sociology occurred in concerns about potential declines in community due
to industrialization and the advent of modernity. Early sociologists spoke
of a shift from Gemeinschaft to Gesellschaft (Tönnies [1887] 1957) or
about the impact of the metropolis on human life (Simmel [1903] 1950).
This decline-of-community thesis has continued to resurface periodically
(e.g., Wirth 1938; Stein 1960; see Lee et al. [1984] or Wellman, Carrington,
and Hall [1988] for reviews of the decline-of-community theses and oppos-
ing viewpoints.)

Many theorists also related declines in community to political outcomes.
This began with Tocqueville ([1835, 1840] 1990) who posited a link be-
tween American democracy and Americans' high rates of joining volun-
tary associations. In later iterations, mass society theorists (e.g., Arendt
1948) claimed that when there is a lack of community, it provides a breed-
ing ground for totalitarianism. So time and time again, individuals raised

the question of a decline in community and its potential political outcomes.

In 1995, Robert Putnam published the thesis that America's social capital is declining. He cited decreasing voter turnout and declining membership in groups such as the PTA and bowling leagues as evidence for a general decline in the ties linking people in the United States to each other and to the political system. The result, Putnam claims, is a massive threat to the successful maintenance of American democracy. His thesis is therefore the latest installment in this long history of speculation on declines in community and its consequences.

What does Putnam mean by a decline in social capital? Social capital is the idea that individuals and groups can gain resources from their connections to one another (and the type of these connections). These resources can be used to produce certain goods. For example, consider a neighborhood with high social capital. In that neighborhood, the neighbors know each other, talk to each other often, and trust each other. In that neighborhood, a mother might feel comfortable letting her child walk alone to a nearby park. In a neighborhood with lower social capital, where the neighbors do not know or trust one another, the mother would either have to walk with her child to the park or hire someone to do it for her.

Putnam is concerned with social capital at a wider level. So, by a decline in America's social capital, he is arguing that certain aspects of social behavior, specifically citizens' rates of joining voluntary associations, citizens' trust in one another, and citizens' rates of voting, are declining. We should therefore see a subsequent decline in certain public goods, such as an efficient democracy.

Like its predecessors, Putnam's thesis sparked a huge debate both in the academic and popular press. For example, *Public Perspective* (1996) and *American Behavioral Scientist* (1997) ran entire issues on social capital and civic decline. Outside of the academic realm, Putnam's thesis appeared in magazines such as *American Prospect* (Putnam 1996) and was critiqued in *Atlantic Monthly* (Lemann 1996). Newspaper commentators used Putnam's thesis to highlight and lament declines in group membership (e.g., Ehrenholt 1995). Putnam was even the subject of profiles in *People* and *U.S. News and World Report*. For all involved, the potential decline of social capital touches on the classic concern that the underpinnings of U.S. democracy may be coming undone.

Despite the amount of interest in a possible decline of U.S. social capital, however, scholars have not reached a consensus on the trend. Putnam has claimed that social capital is in decline while others (e.g., Ladd 1996) have argued that social capital has remained stable over time. I believe this lack of agreement reflects two problems in previous assessments of social capital in the United States. First, there is a large gap between the concept

of social capital and its measurement. Previous studies provide little rationale for how their measures of social capital connect to the theoretical definition of social capital. The problem is compounded by a current lack of consensus on the meaning of the term. The term "social capital" is used in many recent articles but in vastly different ways (e.g., Sanders and Nee 1996; Hagan, MacMillan, and Wheaton 1996; Schiff 1992). The lack of an obvious link between theory and measurement has, in some cases, led to the use of questionable indicators of social capital. For example, voting should be considered an *outcome* of social capital rather than a part of social capital itself.

The second problem with previous assessments of social capital is that they rely on single indicators. Social capital is a general concept, and we should not expect that it can be captured with just one variable. Many different measures can be and have been posited as indicators of social capital. Without strong ties to theory, however, researchers can choose among many pieces of data that provide contrary pictures of the health of social capital in the United States. Also, using measures from a variety of different sources means that assessment is difficult due to incomparability in sampling designs and question wording (Wuthnow 1997). Finally, by using single observed variables, researchers cannot account for measurement error, which we would expect to find in the survey questions used to assess social capital.

In short, to provide evidence for or against recent changes in U.S. social capital, researchers have tracked the mean of a single observed variable over time. If the *relationship* between that variable and the level of actual, unmeasured social capital in the United States has changed over time, however, it could produce a change in the mean of the observed variable, without truly reflecting a change in social capital. For example, survey questions can be interpreted differently by respondents in two time periods. If respondents relax their interpretation of trust between 1975 and 1985, we could see a change in measured trust even if there was no change in the actual level of trust over the 10-year period. Single indicators cannot address this problem.[2]

Previous assessments of social capital are also limited by their singular

---

[2] Other critiques of analyses of social capital note that the *GSS* questions have remained stable while new types of groups have appeared in the United States (Baumgartner and Walker 1988). (The argument is refuted by Smith [1990].) This means that groups, such as environmental groups or other new social movements, which have been growing rapidly over the time period, could be undercounted in the *GSS* measure. Putnam claims that it is forms of group participation requiring minimal effort, such as signing petitions and writing checks, that have increased, thereby accounting for the increase in national environmental organizations, such as the Sierra Club, and organizations like the American Association of Retired Persons (AARP).

focus on change in the *level* of social capital over time. Another important question concerns a possible change in the *dispersion,* or variance, of social capital over time. The issue is similar to income inequality—while the mean level of income in a country might remain the same over time, increasing dispersion would indicate increasing income inequality. To fully understand shifts in social capital over time, therefore, we should investigate possible changes in both the level and dispersion.

The problems with previous assessments of social capital indicate that the current debate over social capital in the United States amounts to a great deal of arguing over selective pieces of information, drawn from different sources and analyzed with weak statistical techniques.[3] Yet, research into the causes of the decline in social capital is moving ahead, under the assumption that it has declined over time (e.g., Brehm and Rahn 1997).

In this article, I address the limitations of previous research by providing a model of social capital that makes explicit links to theory and analyzes multiple indicators of social capital from the same data source. With multiple indicators, I can more adequately gauge the concept of social capital and allow for measurement error. I also illustrate that the relationship between my measures of social capital and the theoretical concept remains stable over time. To begin, I present a theoretical section that defines social capital and distinguishes between the various uses of the term. Next, I present a model of social capital at the national level and discuss the data and variables I use to estimate it. I then test whether my chosen measures have a stable relationship with social capital over time and include a test for change in the variance of social capital over time. In the final sections, I estimate change in the level of social capital over a 20-year period, considering both linear and nonlinear trends.

WHAT IS SOCIAL CAPITAL?

The idea of social capital can be placed in a historical series of ideas on different forms of capital. Originally, the concept of physical capital was introduced to explain the ways that physical implements, such as tools or machines, could facilitate economic production. Then Becker (1964),

---

[3] Similar concerns surfaced in the past under different names, e.g., "the confidence gap" (Lipset and Schneider 1983), "the crisis of legitimacy" (Lehman 1987), "mass society" (Halebsky 1976). In addition, many theorists viewed aspects of social capital, such as voluntary associations or trust, as central to a successful social life. A list of social theorists who commented on the importance of trust or associations to social life would read like a "who's who" of social theory—Marx, Durkheim, Weber, Locke, Simmel. For a discussion of the classical roots of social capital, see Portes and Sensenbrenner (1993).

building on Schultz (1961), presented the notion of human capital and argued that individuals, through education or job training, can hold *within themselves* the ability to facilitate production.[4] The newer concept of social capital acknowledges that certain social relations (e.g., dense networks, norms of reciprocity) can also facilitate production. In addition, with the introduction of social capital, researchers began to speak of efficiency gains in noneconomic goods.

The concept of social capital was introduced by two major social scientists—Bourdieu (1983) and Coleman (1988, 1990).[5] A number of other authors, from a variety of fields, also used the term (e.g., Hanifan 1920; Jacobs 1961; Loury 1977), but it remained obscure until Bourdieu and Coleman popularized it.[6] Bourdieu (1983, p. 248) provides a concise definition: "Social capital is the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition—or in other words, to membership in a group." Social capital requires more than just a network of ties, however. Bourdieu notes that social capital also involves "transforming contingent relations, such as those of neighborhood, the workplace, or even kinship, into relationships that are at once necessary and elective, implying durable obligations subjectively felt (feelings of gratitude, respect, friendship, etc.)" (1983, pp. 249–50). Thus, network ties must also be of a particular type—trusting and positive.

For Coleman (1988, p. S98), the concept of social capital illustrates how the social structure of a group can function as a resource for the individuals of that group. He claims that social capital "inheres in the structure of relations between actors and among actors." Thus, social capital is not lodged in individuals themselves, although they can make use of it to facilitate the production of individual or collective ends. Coleman sees the existence of social capital in trust, information, norms and effective sanctions, authority relations, and the extent of obligations in a group. Each is a feature of the social structure that also provides social capital as a resource for the individuals of the group.[7]

---

[4] According to Schultz, the notion of human beings as capital appears in the work of several classical theorists, including Adam Smith.

[5] Woolcock (1998) provides an exceptionally extensive review of the use of social capital.

[6] Hanifan's (1920) early reference to social capital means that the concept, although forgotten for over 50 years, actually predates the notion of human capital. I thank Robert Putnam and Brad Clarke for pointing out this citation on SOCNET, the social networks listserv.

[7] Becker (1996) has recently expanded his treatment of human capital to include various other forms of capital such as personal capital, imagination capital, and social capital. He sees an individual's social capital as part of his total stock of human capital, where social capital is the relevant past actions by peers and others in an

Coleman (1988) uses an example of social capital that is extremely useful in understanding the concept.

> Wholesale diamond markets exhibit a property that to an outsider is remarkable. In the process of negotiating a sale, a merchant will hand over to another merchant a bag of stones for the latter to examine in private at his leisure, with no formal insurance that the latter will not substitute one or more inferior stones or a paste replica The merchandise may be worth thousands, or hundreds of thousands, of dollars. Such free exchange of stones for inspection is important to the functioning of this market. In its absence, the market would operate in a much more cumbersome, much less efficient manner. (P. S98)

Coleman observes that within this wholesale diamond market, individuals have many close ties through family, community, and religious affiliation. There is an "extra layer" over their business transactions that allows trustworthiness to be taken for granted. This "social capital" allows for efficient economic transactions—the diamond merchants can avoid expensive bonding and insurance devices.

To concisely represent and summarize previous theoretical work on the subject, I suggest the following definition of social capital. Social capital involves two components:

1. *Objective associations between individuals.*—There must be an objective network structure linking individuals. This component indicates that individuals are tied to each other in social space.

2. *A subjective type of tie.*—The ties between individuals must be of a particular type—reciprocal, trusting, and involving positive emotion.[8]

When social capital is present, it increases the capacity for action and facilitates the production of some good. When active, it facilitates various ends for the members of a group and for the group as a whole. Social capital could, however, remain latent within the group and be viewed as potential energy.

The goods produced by social capital can occur at different levels of the social structure.[9] For example, we can imagine the production of an

---

individual's social network that affect current or future utilities. This formulation is complementary to the one presented in Coleman and Bourdieu but stresses the importance of others' choices in the total amount of social capital available to any individual or group.

[8] This two-component definition of social capital reflects the traditional division in social theory between structure and content (Simmel 1971). Seen another way, social capital has both a quantitative and qualitative dimension.

[9] This point has not been recognized in previous research. Instead, researchers have argued that social capital at the level they consider is the only "real" social capital. However, social capital, as a general concept, can be measured at multiple levels, just as Becker (1964) or Schultz (1961) describe human capital at both the individual and aggregate level.

individual, private-level good—a mother asks a friend to baby-sit rather than hiring a baby-sitter. In this case, social capital is an individual, private good that, like human capital, can be used for economic gain or another private outcome such as educational attainment. Saunders and Nee (1996) use social capital in this way when they employ family-level variables and consider their effect on individual family self-employment. Other researchers consider the impact of migration (as a breakdown of social capital for an individual family) on an individual child's educational attainment (Hagan et al. 1996; Smith, Beaulieu, and Seraphine 1995). Social capital, as ties between individuals and the context of those ties, in this case produces an individual-level good—one person's ties produce a good that is used only by them. Of course, if a mother asks a friend to baby-sit, she is actually increasing the social capital between them, since she is incurring an obligation. The good produced in this single instance, however, remains primarily hers.

We can also consider the production of group-level goods. The diamond merchants provide an excellent example of this level—any of the merchants can draw upon the accumulated social capital of the whole group. Or, the social capital in such a group can benefit *all* members at the same time if they collectively pursue a public good. Individual and group-level social capital are linked. An individual may hold a number of obligations that can benefit him or her personally; at the same time, as part of a larger system of obligations, these obligations can contribute to the social capital of a group. To take an example from Portes and Sensenbrenner (1993), a Latino police officer fighting a court battle after shooting two black cyclists was able to use pleas for help on Hispanic radio stations to raise money. The tight ties and positive feelings in that community were used in his case to benefit a single individual. The social capital in that community could have benefited any individual member, however. Studies that consider the positive impact of ethnic subgroups on all their members are examples of the production of group-level goods (Zhou and Bankstron 1994). Or, consider how a family's commitment to education and commitment of other families in their school combine to determine the success of an individual child.[10]

---

[10] It is also useful to consider a case where the production of an individual-level good is tightly tied to an even more general level of social capital. Consider a game representing trust in a noncontract exchange situation, such as an illegal drug market (see Dasgupta 1988, pp. 61–63). In this situation, a buyer has the choice between trying to buy a high-quality drug (placing trust) and not trying to buy the drug (not placing trust). If the buyer chooses not to buy the drug, the game is over. If the buyer decides to buy, however, then the seller has a choice between giving a drug of high quality (keeping trust) or giving a drug of lower quality (breaking trust). The buyer would prefer to have the high-quality drug, but would also prefer to have no drug over receiving a poor-quality drug. In this game, population-level trust determines whether

## Associations between Individuals

| | | high | low |
|---|---|---|---|
| **Trust, Reciprocity, & Positive Emotions** | high | social capital | |
| | low | | |

FIG. 1.—Social capital

Social capital can also be considered to produce goods at the community level (between many groups).[11] While some theories speak of social capital within a single group of individuals, like the diamond traders, other theories, such as Putnam (1995) or Fukuyama (1995), utilize social capital as a macrosociological phenomenon, or a feature of a community. In their theories, entire nations can hold differing levels of social capital, which in turn affect the chances for democracy or industrialization.

Social capital, as I have defined it, involves the mutual occurrence of two components, which in turn produces a capacity for action. Another way to view social capital is with a 2 × 2 table, as in figure 1.

Social capital exists in the top left corner of the table. Viewing social capital in terms of a 2 × 2 table is a useful explanatory device. For example, I can place the diamond market example within the 2 × 2 table. Figure 2 illustrates that it is only in the presence of both high associations and high trust that we see the economic efficiency gains from the trusting exchange of diamonds. Otherwise, it takes the addition of external agents

---

the individuals will exchange. That is, if the buyer believes the trustworthiness of the average person is high, he or she will choose to make the exchange. If we restrict ourselves to the creation of trust between those two individuals in the absence of others, however, individual-level social capital cannot be an explanation (because social capital itself involves trust). Instead, to explain this exchange, the different levels of social capital become very important. Also, this example illustrates the temporal element to social capital. If social capital existed between the two individuals in a previous time point, then it could explain the creation or continuance of trust in the present time period, regardless of the population level of trust.

[11] A similar distinction between these three levels (individual, within-group, and between-group) is found in human capital research, where individuals personally hold human capital such as education or job training, yet researchers also discuss a company's attempt to raise levels of human capital, or review the overall level of human capital in the United States (e.g., Becker 1964, pp. 23–24).

Associations between individuals

| | | high | low |
|---|---|---|---|
| Trust, Reciprocity, & Positive Emotions | high | trusting exchange of diamonds | trust or goodwill present but problem with interaction (e.g., kin in different cities or structural barriers to interaction, such as so much crime that jewels cannot be carried around) |
| | low | a more typical neoclassical business transaction; with little trust, the diamond merchants would need to use the legal system when exchanging diamonds | |

FIG. 2.—The diamond market example

for action to occur. In the high trust/low association cell, intermediaries in the form of go-betweens or transporters would be necessary for action. In the low trust/high association cell, a third party is needed to insure or enforce the transaction.

One major problem with previous theoretical work on social capital is that researchers have assumed that its effects will always be positive. However, acknowledging that social capital can exist at different levels can help us understand that social capital need not always imply positive effects for all members of a community. That is, social capital *within a single group* need not be positively related to social capital *at the community level.* While social capital within a particular group may be expected to have positive effects for the members *of that group,* this need not "spill over" into positive gains in social capital for the community. For example, a militia or ethnic separatist group might have high social capital within its individual group but reduce social capital in the larger community by either having no ties with others outside of the group or in reducing the overall level of trust in those outside of the group. Granovetter (1973) speculates on the importance of weak ties to community organization— in a similar manner, ties between groups could be necessary to increase community-level social capital.

Not only can social capital within a single group potentially reduce social capital between groups, but high within-group social capital could have *negative* effects for members of the community as a whole. That is, the potential energy created by an individual group could be used for nefarious purposes. To take an example from Gambetta (1988, p. 214),

**Within-Group Trust and Associations**

|  |  | high | low |
|---|---|---|---|
| **Between-Group Trust and Associations** | high | community-level social capital increased; expected positive benefits for the community | should not see many examples of this cell |
|  | low | community-level social capital decreased; potentially negative effects on the community | low social capital overall |

FIG. 3.—Community-level social capital

there are instances of high social capital, like those among robbers and murderers, that we would want to reduce for the sake of the community as a whole. This is a point that has been made in research on civil society (Nelson 1994, p. 150; Fatton 1995, p. 72) and trust (Gambetta 1988, p. 214) but is only recently acknowledged with regard to social capital (e.g., Portes and Sensenbrenner 1993; Portes 1998; Foley and Edwards 1997).

We would expect to see decreased social capital or negative effects at the level of the community when there is low between-group trust and networks but high within-group trust and networks. In this case, organized interest groups could effectively decrease trust within the entire community and possibly result in outcomes that do not benefit the community as a whole, or every member of the community.

Positive, community-level social capital would be expected to occur when there are positive, trusting ties between individuals in different groups (crosscutting ties). As before, this situation can be illustrated with a 2 × 2 table (see fig. 3).[12]

MORE ON SOCIAL CAPITAL: ITS COMPONENTS AND EFFECTS

By investigating the potential decline of social capital in the United States, this article considers aggregate, positive social capital (the top left-hand cell of fig. 3). In the preceding section, I illustrated that social capital involves two components: trust and associations. In this section, I consider

[12] I thank Jonathan Hartlyn for suggesting a 2 × 2 representation of this point.

each component in depth and discuss how it can be measured in an aggregate manner. I also illustrate, with an extended example, how aggregate measures of social capital can have aggregate-level effects.

Trust

The first component of social capital relates to the types of ties between individuals, where the presence of positive ties is essential. For the measurement of aggregated social capital, I focus on trust. An emphasis on trust, over other types of ties, is prevalent in the literature (e.g., Putnam 1995; Brehm and Rahn 1997), as it is difficult to measure positive emotions or the extent of reciprocity at the national level. Most theory is driven by discussions of trust, and trust is highly associated with generalized reciprocity, so trust remains a good proxy for positive, reciprocal ties in general.

Barber (1983, p. 165) defines trust as "socially learned and socially confirmed expectations that people have of each other, of the organizations and institutions in which they live, and of the natural and moral social orders, that set the fundamental understandings for their lives." Trust can occur in at least three levels of the social structure: in the isolated dyad, between individuals in the presence of third parties, and between an individual and a collection of individuals, such as an organization or an institution.

In measuring aggregate trust, I focus on perceptions of trustworthiness rather than the actual placement of trust. At aggregate levels, it is often not reasonable to consider the trustor as having full choice in the placement of trust. For example, as Luhmann (1979) explains, to not trust in its broadest sense would prevent an individual from rising in the morning. He calls this type of trust "confidence" (Luhmann 1988). Also, with many institutions, the trustor does not have any choice about whether to place trust. For example, it is illegal for a citizen to choose not to place trust in the U.S. government by not paying taxes. This does not mean, however, that individuals do not have an opinion about the trustworthiness of an institution.

Theorists make a distinction between trust in specific individuals and trust in more abstract people or systems (e.g., Giddens 1990). Abstract trust entails less awareness of risk (Luhmann 1988) and less awareness of the person being trusted. So, while actors can make very specific determinations of the trustworthiness of certain individuals (based on information about that individual's history, motivation, competence, etc.), they may also hold opinions about the trustworthiness of more generalized others. For example, individuals may have an opinion about the trustworthiness of the "average" person. In measuring aggregate trust, I focus on individu-

als' estimates of the trustworthiness of generalized others, or abstract trust. While trust in specific others may be important at more microlevels of social capital, generalized trust is the important feature of national-level social capital.[13]

Individuals can also hold opinions about the trustworthiness of abstract systems, such as institutions, which are aggregations of individuals embedded in particular social structures. Based on generalized estimates of the technical competence and moral obligations of individuals in an institution, as well as estimates of the sanctions inherent in the social structure of that institution, a person in the United States can develop institution-specific levels of trust. This is similar to Giddens's (1990) notion of trust in expert systems, where an actor may not know the person who built their car or their house, but they trust the system of accreditation, regulation, and monitoring in which the person is embedded. Previous assessments of social capital have concentrated only on trust in governmental institutions. This was mainly due to the conflation of social capital with its outcomes (civic engagement), however. Instead, trust in many aggregate institutions is necessary for an assessment of national-level social capital.

There are a few other types of trust that fall between trust in a specific individual and trust in generalized others or institutions. For example, when a potential trustor is embedded in a group, he or she may assign the other members of that group a level of trustworthiness that is higher than the trustworthiness accorded to the average person, due to the presence of norms and sanctions against those who break trust. In the same manner, a potential trustor might identify certain categorical groups in society to receive more or less trust than the average person. For example, many people in the United States hold notions of the typical criminal, which could cause them to trust young, black men less than other categorical groups (Reiman 1990). Or, some people in the United States may believe that groups such as lawyers or politicians have low feelings of moral obligation and consequently downrate their level of trustworthiness as an entire group. These distinctions between in-group and out-group trust are important for distinguishing between the creation of positive, community-level social capital and some alternative negative outcomes (i.e., figure 3).

## Associations

The second component of social capital reflects the objective ties between individuals—their associations with each other. Associations between in-

---

[13] Dasgupta (1988) in a game-theoretic model of trust, discusses a population's reputation for honesty. In his model, while a potential trustee may be either honest or dishonest, a trustor makes use of the entire population's reputation for trustworthiness in making a decision whether to place trust.

dividuals fall into two types. Individuals can be informally connected to others through friendship choices and other types of network ties, or individuals can be connected to others through formal group memberships. First, individuals have relationships with other individuals. That is, individuals have a social network. Relationships (ties) can be of many different types, including friendship or other emotional ties; transfers of material resources, or exchange relationships; proximity in space, such as neighbors or office mates; and kinship relations (Wasserman and Faust 1994). They can also be directional or nondirectional, valued or dichotomous, and uniplex or multiplex (see Wasserman and Faust [1994, chaps. 1–2] for a discussion of network terminology). Each of these features of the association has implications for the total stock of social capital. At its base, however, is the argument that an individual's informal friendships with old schoolmates, fellow workers, or the friend of a friend can create social capital through increased communication, information diffusion, and social support.

Besides informal ties to others, individuals can be tied to other individuals through formal membership in voluntary associations (an association or affiliation tie). With association ties, individuals are linked through their joint presence at an association event, or through their joint membership in an association (Breiger 1974). To survive over time, voluntary associations must recruit and maintain members. This is one way to distinguish formal associations from informal networks: informal friendship networks are defined *by* the ties between individuals, but formal associations survive beyond any particular member or internal social network. In voluntary associations, in addition to the benefits of network ties, members access and create additional group-level benefits. It is this membership in *groups*, not simply the ties between individuals, that provides further resources to solve collective problems and pursue specific goals in a large society. Much theory and research has illustrated this potential in groups (e.g., Lipset, Trow, and Coleman 1956).

For small groups, the social network of individuals and all their group memberships can often be defined exactly through a complete enumeration of ties (e.g., Moody and Bearman 1997). Network characteristics, such as density and reach, on a variety of ties, such as obligations and friendships, could accurately reflect the structure of the overall network of individuals and their ties to groups for the measurement of social capital.

In larger groups like nations, a complete enumeration is impossible. In those cases, a random sample of the ties between individuals provides only an estimate of the density of the aggregate network.[14] National surveys ask

---

[14] This means that there is a high possibility of inaccurate measurement, which should be accounted for in our models. The model I will present does account for measure-

questions about the number and type of group memberships, which can be used to estimate the general level of associations through groups in the nation as a whole. In addition, more traditional network questions in national surveys allow an estimation of the ties between individuals at the larger level.[15]

It is also desirable to consider the ties between individuals across various groups (crosscutting ties). As discussed above, it is these types of ties that can translate into positive social capital for the community as a whole. Unfortunately, without a census of associations and the ties between them, it is difficult to measure these crosscutting ties at the national level. The best we can do is assume that when individuals have multiple association memberships, it indicates some connection between the associations through individuals (see Breiger 1974). It is essential to better measure and test connections between associations in future research.

## Other Possible Components of Social Capital

Political participation and volunteering are not included in my model of social capital, although some have treated them as indicators (e.g., Putnam 1995; Ladd 1996; Wuthnow 1997). Social capital, as originally theorized, does not include specific actions of individuals, such as voting or volunteering—these are outcomes that we would expect to be *facilitated* by high levels of social capital. Once outcomes are separated from social capital, we can test whether declining levels of social capital have detrimental effects on other variables such as voting.

## Efficiency and Productivity Gains at the Aggregate Level

By investigating the potential decline of social capital in the United States, this article considers aggregate-level social capital. The previous two sections illustrated how the two components of social capital can be measured in an aggregate manner. The final portion of my definition of social capital implies an increased capacity for action. At the aggregate level, this should imply aggregate-level gains in productivity or efficiency.

In general discussions, trust is described as a lubricant that eliminates the need for third-party insurers or enforcers. At the same time, group memberships, especially crosscutting ones, are expected to increase com-

---

ment error, and I believe it is the first model of social capital to take measurement error into account.

[15] More complex sampling designs are also possible. For example, McPherson's (1982) hypernetwork sampling design links a representative sample of organizations to a representative sample of individuals

munication and information flows. Both these effects of social capital, at the aggregate level, can increase efficiency and productivity in any number of aggregate outcomes. For example, Fukuyama (1995) outlines the benefits for national economic efficiency (see also, Granovetter 1985), Coleman (1988) focuses on reductions in high school dropout rates, and Putnam (1993) argues for increases in government efficiency. To illustrate the relationship between aggregate-level social capital and aggregate-level public goods, I will focus on the maintenance of democracy—an important aggregate-level public good and the one most often considered "in danger" by social capital researchers.[16] I provide one link between social capital and democracy for each component, although others are possible. A more extensive discussion of the relationship between democracy and social capital can be found in Paxton (1998).

Consider high national levels of trust in generalized, or random, others. Such trust is extremely important in a democratic system because individuals must be willing to place political power in the hands of "the people." With low levels of observed trust, individual citizens would be unwilling to relinquish political power to those with opposing viewpoints, even for a short time.[17] In democracies, there must be competition between groups, and power can shift from one group to another. Without trust, individuals would not believe that others would follow the "rules of the game" while holding power, and so would be unwilling to give up power themselves, even for a short period. A lack of trust could therefore seriously undermine the continuance of democracy, because there would be no way to successfully maneuver turnovers of power.[18] High aggregate levels of trust, on the other hand, would help ensure efficient, regular turnovers of power. Without protracted succession battles, the health of a democracy and its ability to productively address national issues are enhanced.

As for the other component of social capital, associations, the overall connectedness of a population increases information flows. Increased information flows aid in the maintenance of democracy by ensuring that political participation is tolerant, moderate, and publicly oriented. Tocqueville (1990) argued that as individuals participate in associations,

[16] Transitions to democracy require separate arguments, see Paxton (1998).

[17] Trust in one's particular group, or only a segment of the population, would be inadequate. In a pluralist system, individuals must display *generalized* trust, since at any time, any group could obtain power.

[18] Trust in institutions is also important for the maintenance of democracy through its relationship to democratic legitimacy Even if individuals have low generalized trust in individuals, they may trust "the system" enough to allow politically objectionable groups to hold power for a time. However, some level of trust in individuals is still necessary to believe that others will uphold the system. Democratic legitimacy is discussed extensively elsewhere (e.g., Lipset and Schneider 1983; Lehman 1987).

they see others who are also participating in associations and notice that their interests coincide at a greater level. This learned similarity causes an individual to develop an "enlightened self-interest," which moves beyond individual self-interest to a consideration of the public good, the promotion of a common identity, and a sense of shared responsibility. In addition, with more aggregate-level association memberships, new ideas and opinions are more quickly disseminated throughout the population, yet, extremist ideas are more easily challenged, as they have less chance of remaining isolated. An "enlightened self-interest" and the checks on extremist ideas mean that individuals should have a wider interest in mind when participating in politics, thereby changing the type (i.e., more tolerant, moderate, etc.) of their democratic participation.

Empirical evidence supports the view that social capital is related to the maintenance of democracy. First, a massive amount of research (e.g., Verba, Schlozman, and Brady 1995; Verba, Nie, and Kim 1978; Wolfinger and Rosenstone 1980) has shown that membership in voluntary associations stimulates political participation. A related line of research illustrates that ties to the community increase the chances of individual political participation as well (e.g., Guest and Orpesa 1986). Although there is some debate on the issue, many see extensive political participation as essential for the health of a democracy (Mill [1831] 1975; Tocqueville 1990; Arendt 1948). Second, Putnam's (1993) research in Italy considered the relationship between social capital and democratic governmental performance. He found a measure of "civic community" (which included association memberships) to be highly correlated with democratic institutional performance ($r = 0.92$). (He also noted that the correlation remained significant with a control for economic development.) Third, a cross-national analysis of democracy and social capital reveals a similar positive relationship (Paxton 1998). The correlation coefficient between measures of trust and democracy across 45 countries is $r = 0.44$ ($p < .003$),[19] and remains if industrialization is controlled (partial $r = 0.42$ [$p < .005$]). And, the partial correlation (controlling for industrialization) between democracy and a count of international nongovernmental organizations in over 150 countries is 0.35 ($p < .0001$).[20]

---

[19] I report the correlation with two extreme outliers (China and Nigeria) removed from the sample.

[20] While I have focused on only one outcome, similar arguments could be presented for other public goods. The lubrication, predictability, connectedness, and communication that come from high levels of social capital increase efficiency in economic transactions, scientific endeavors, and the political process. To briefly outline another example, Becker (1964) explains how increases in human capital aid economic development through the growth of scientific and technical knowledge (which raises the productivity of labor and other production inputs). In a similar manner, increased associations and trust would help quickly disseminate technical or scientific knowl-

To recapitulate, both components of social capital, trust and associations, can be measured at the aggregate level. Individuals can make assessments of the trustworthiness of generalized others, as well as people associated with specific national institutions. Individuals can have objective ties to others through their membership in voluntary associations or outside of such groups, in their ties to neighbors or friends. At the national level, high levels of trust and association memberships should enhance public goods such as the maintenance of a healthy democracy.[21] In the next section, I propose a model of social capital that reflects both components of social capital at the national level.

MODEL

Assessing a possible decline in social capital requires a model of social capital that incorporates multiple indicators over time. I use data from the *General Social Surveys* (Davis and Smith 1994), or *GSS*, because it contains multiple indicators for both components of social capital over a 20-year period (1975–94). Specifically, my model of social capital, consisting of 12 indicators and three latent variables, is available for nine years: 1975, 1978, 1983, 1986, 1988, 1990, 1991, 1993, and 1994, where each year samples a different set of individuals.

To conform to my theoretical model, I separate the measurement of social capital into two pieces. One component of the model measures an individual's subjective trust toward others in the community. The second component measures the objective extent of an individual's associations, or ties to the community. For ease in the presentation of my model and the results, I separate the two components and discuss each in turn.[22] In a later section, I explore the combination of trust and associations as a measure of social capital.

---

edge. (The importance of associations for this purpose is obvious See Hardwig [1991] for a discussion of the importance of trust in science.) While much technical or scientific dissemination would take place through professional associations, the spread of knowledge about the Internet illustrates that more general voluntary associations or neighborhood/friendship ties play an important role as well.

[21] Although I spend little time discussing other types of capital, all the types of capital are complementary in their returns on efficiency. For example, human capital and social capital can magnify each other's effects Human capital can complement social capital: a group can make use of the knowledge base of its members when attempting to achieve a collective good. Or, social capital can complement human capital. ties to other skilled individuals can augment an individual's own human capital.

[22] It is not necessary to provide a hypothesis about the effect of one component on the other during measurement. Others have hypothesized some effects (e.g., Brehm and Rahn 1997), but the incorporation of such effects is unnecessary in the present analysis.
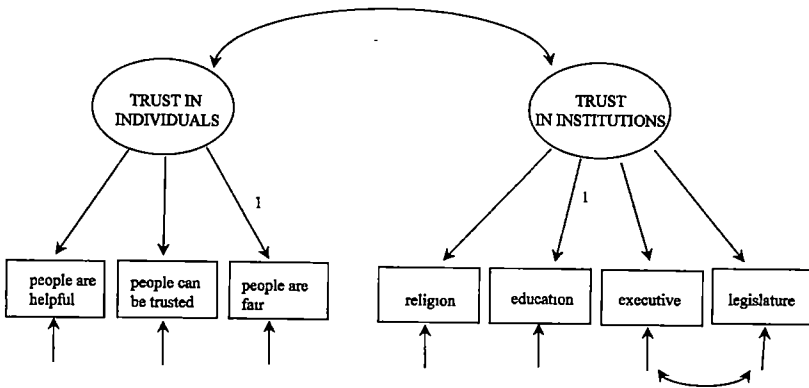
FIG. 4.—Model of the trust component of social capital

Trust

Figure 4 presents my model of the first component of social capital—
trust.[23] I model trust as a confirmatory factor analysis with two dimen-
sions: an individual's general trust in others, and an individual's trust in
institutions.[24] The first dimension, trust in individuals, has three dichoto-
mous observed indicators: "Would you say that most of the time people
try to be helpful, or that they are mostly just looking out for themselves?"
(HELPFUL), "Do you think most people would try to take advantage of
you if they got a chance, or would they try to be fair?" (FAIR), and "Gener-
ally speaking, would you say that most people can be trusted or that you
can't be too careful in dealing with people?" (TRUST). These questions
were first formulated by Rosenberg (1956). Although only one of the vari-
ables uses the word "trust," all three reflect the trustworthiness or integrity

[23] Path diagrams like fig. 4 represent relations between observed (measured) and unob-
served (latent) variables. Latent variables are enclosed in ovals, while observed vari-
ables are represented with boxes. Straight arrows indicate a causal relationship be-
tween two variables, while curved two-headed arrows indicate a covariance between
two variables that is unexplained in the model. Measurement error is indicated by
$\delta$'s, and errors in equations are indicated by $\zeta$'s. Allowing measurement error ac-
knowledges that the variables are not perfect measures of their underlying concepts.
Also, having multiple measures of the same latent variable means that more informa-
tion is available about the concept of interest

[24] To measure the distribution of general trust across the United States, we would also
want to measure the differential trust assigned to various groups. However, measures
of differential group trustworthiness rarely exist in social surveys (a notable exception
is the PEW Research Center study on Philadelphia). The *GSS* does not contain ques-
tions related to trust of different groups, so future research should consider this as
well.
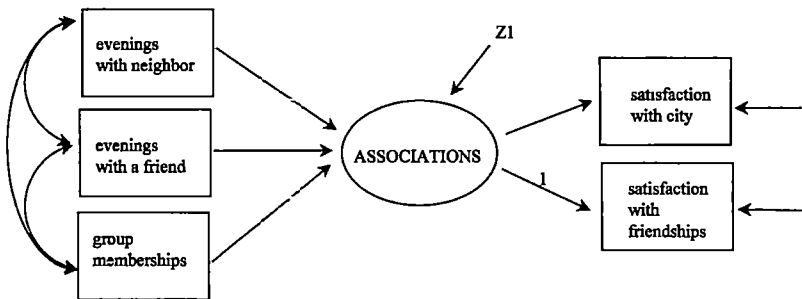
FIG. 5.—Model of the associations component of social capital

of others. Each question queries about feelings of moral obligation, which Barber (1983) theorizes as essential to trust. I model the three questions so that an increase in an individual's feelings about the trustworthiness or moral obligation of others is expected to simultaneously increase answers to all three *GSS* questions.

Trust in institutions considers three general institutions: organized religion (RELIGION); the education system (EDUCATION); and the government, where the government is broken into the executive branch (EXECUTIVE) of the federal government and the Congress (LEGISLATURE). Each of these variables is created from this question: "I am going to name some institutions in this country. As far as the people running those institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?"[25] While others have referred to these questions as "confidence in institutions," I refer to them as "trust" in institutions to emphasize that they represent just another dimension of overall trust.[26]

### Associations

Figure 5 presents my model of individuals' associations, which is different from a typical confirmatory factor analysis. In this model, I expect three

[25] I do not include trust in other institutions such as the military, medicine, and banks, even though the *GSS* asks about them. In deciding which of the 13 available institutions to include, I attempted to choose those that would provide the best information about generalized trust in institutions. To pick the institutions with the most relevance to individuals, I used those that most often had chapters devoted to them in a sample of introductory sociology textbooks.

[26] Such a strategy is supported by the 1978 *GSS*, which asked respondents what the word "confidence" meant to them. Smith (1981, p. 169) reports, "The overall favorite choice was that confidence in the people running institutions means trusting them. Almost 35% mentioned trust in their responses." Other often-mentioned responses were "capability," "believe in," and "faith."

indicators to increase an individual's unobserved level of association: how often a respondent spends a social evening with someone in their neighborhood (EVENINGS WITH NEIGHBOR), how often a respondent spends an evening with friends who live outside the neighborhood (EVENINGS WITH A FRIEND), and the total number of memberships the individual has in voluntary organizations (GROUP MEMBERSHIPS). Together, these three variables measure the objective extent of an individual's associations in the community: both associations with other individuals and associations through groups. As each of these variables increases, I expect an individual's general level of association to increase. So, I place the variables as *causes* of a general level of association, rather than as *effects*, as would be the case in a typical confirmatory factor analysis.

In turn, I expect two indicators to be influenced by an increase in an individual's general level of association: the amount of satisfaction a respondent receives from the city or place in which they live (SATISFACTION WITH CITY), and the satisfaction the respondent receives from friendships (SATISFACTION WITH FRIENDSHIPS).[27] Many studies have linked social support to happiness (see Argyle 1992). So, individuals who are more embedded in their community, through their ties to others or their memberships in groups, are expected to rate their satisfaction with city and friends higher. These two variables are measured on a seven-point scale ranging from no satisfaction to a very great deal.

RESULTS

There are a number of steps to estimating the possible decline of social capital in the United States. (1) Since there are many possible specifications of social capital, I must estimate both components of social capital and establish that my model of social capital fits the data well in every year. (2) Once I have illustrated that my model fits the data well, then I must demonstrate that the relationship between the general level of social capital and my observed measures remains stable over the 20-year period. As noted above, by charting means over time, others have made the assumption that the relationship between the indicators and the latent variable of social capital has not changed over time. This need not be true, and if the relationship has changed over time, we could see a decline in

[27] Similar questions were asked about family interaction and satisfaction. I do not use those variables in the model, however, as family interaction is widely seen as qualitatively different than community interaction. In fact, some scholars have posited that excessive family interaction is detrimental to community-level interaction (e.g., Banfield 1958).

a measure of social capital *even if* there was no change in the latent level of social capital over the time period. I must therefore establish that my model's parameters have remained stable over time or I cannot say anything about a change in general social capital.[28] (3) Once I have established that my model fits well and that its parameters are stable over time, I can finally move to the question of a decline in social capital. In doing so, I consider both linear and nonlinear trends.[29]

## The Fit of the Model in Each of the Nine Years

The first question I must answer in analyzing the data is whether my model fits the data well in every year. Fit statistics for both components of the model in each year appear in table 1. I provide fit statistics from various families (Tanaka 1993), which together give a comprehensive view of how closely the model represents the data in each year. The top half of table 1 indicates that the trust component of the model fits well in each year. The only indication of a poor fit is the significant chi-square test statistic in four of the nine years. However, the chi-square test is sensitive to sample size and the distribution of the observed variables. Because the chi-square test is a measure of "perfect" fit, any slight deviation from a perfect fit could induce a significant test, especially if the model has a great deal of power (e.g., the sample size is large). Other fit statistics are less sensitive to sample sizes, and these, the root mean square error of approximation (RMSEA) (Steiger and Lind 1980) and incremental fit index (IFI) (Bollen 1989*b*), indicate an excellent fit in each year. The ad-

[28] This section serves two additional purposes First, by testing whether the parameters of the model remain stable over time, I can test for a change in the variance (or dispersion) of social capital over time. Second, once I provide evidence that my parameters are stable over time, I can pool the years together. One cannot pool data sets without first establishing that the parameters across the groups are the same.

[29] In analyzing the data, each component of the model required different corrective procedures. The trust component of the model includes categorical endogenous variables (dichotomies and trichotomies), so I used PRELIS to calculate the polychoric correlation matrix, which was then estimated in LISREL with weighted least squares (Jöreskog and Sörbom 1993). The polychoric correlation matrix is an estimate of the correlation between the two continuous variables that underlie the categorical variables, based on an assumption of normality. Used in conjunction with the asymptotic covariance matrix, this procedure produces consistent estimates of the parameters and unbiased standard errors. The associations component of the model had a large number of missing values To take advantage of cases with incomplete data, I performed a maximum-likelihood estimation of the model with missing values in AMOS (Arbuckle 1995). In addition, I estimated the associations component with listwise deletion of missing values so that the results could be compared for extreme discrepancies.

TABLE 1

THE FIT OF THE MODELS OF TRUST AND ASSOCIATIONS IN EACH YEAR

| Year | $N$ | $\chi^2$ | df | $p$-value | AGFI | RMSEA | IFI |
|---|---|---|---|---|---|---|---|
| Trust: | | | | | | | |
| 1975 . ...... | 1,150 | 20 75 | 12 | 054 | 99 | .025 | .99 |
| 1978 ........ | 1,245 | 18 2 | 12 | 11 | .99 | .02 | 1 |
| 1983 ... ... | 660 | 18 44 | 12 | .1 | .99 | 029 | .99 |
| 1986 .. .. | 1,251 | 23.2 | 12 | .026 | .99 | .027 | .99 |
| 1988 . ... .. | 813 | 15.88 | 12 | 197 | 99 | .02 | .99 |
| 1990 ........ | 722 | 26.66 | 12 | 009 | .98 | .04 | .98 |
| 1991 ........ | 824 | 18 47 | 12 | .1 | .99 | .026 | .99 |
| 1993 .. ... | 850 | 38.58 | 12 | .0001 | .98 | .05 | .96 |
| 1994 ........ | 1,633 | 21.76 | 12 | .04 | 99 | .022 | .99 |
| Associations. | | | | | | | |
| 1975 ........ | 1,490 | 19 74 | 2 | .00005 | .96 | 077 | .95 |
| 1978 . ... | 1,532 | 24.65 | 2 | 0006 | .95 | .086 | .93 |
| 1983 ... ... | 1,599 | 15.97 | 2 | .004 | .97 | .066 | .96 |
| 1986 ........ | 1,469 | 8.36 | 2 | .015 | .98 | .047 | .98 |
| 1988 ..... | 1,480 | 3.83 | 2 | 147 | .97 | .025 | .99 |
| 1990 .. ... | 1,369 | 5 105 | 2 | 078 | .97 | .034 | .98 |
| 1991 ........ | 1,514 | 4 086 | 2 | .13 | .97 | 026 | .99 |
| 1993 ........ | 1,602 | 9.058 | 2 | .011 | .98 | 047 | .97 |
| 1994 . ...... | 2,213 | 9.647 | 4 | .047 | 94 | .025 | .96 |

NOTE.—Trust model results are figured with listwise polychorics (PRELIS/LISREL), association model results are figured with MLE missing values (AMOS)

justed goodness-of-fit index (AGFI) (Jöreskog and Sörbom 1986) also indicates an excellent fit.[30]

The associations component of the model also illustrates a good fit in each year.[31] Although the chi-square test statistic is significant in the earlier years, this could be a function of the larger sample size for cases with complete data in those years. In this component of the model, the RMSEA is occasionally a bit high, although still within acceptable range (Browne and Cudeck 1993, p. 144). The other two measures of fit indicate a good

[30] The closer the AGFI and the IFI are to 1.0, the better the fit of a model. In contrast, the closer the RMSEA to 0, the better the fit of the model.

[31] The AGFI fit statistic is not available in the AMOS missing value routine, so I present the AGFI from the listwise deleted missing value estimation in CALIS. The sample sizes for the listwise deleted estimation are (1975–94): 1474, 1511, 1581, 1444, 497, 452, 520, 526, and 266.

to excellent fit in all years.[32] Therefore, the measurement models for trust and associations are consistent with the data.

### The Stability of the Parameter Values over the Time Period

As discussed above, the next step is to test whether the parameters of the model (the factor loadings, structural parameters, variance of the latent variables, and the variance of the errors) remain the same over time. If they do not change, then I can track the level of social capital over time. Otherwise, any estimated change in the level of a latent variable could be due to a difference in parameter values across time (see Bollen [1989a, pp. 355–60] for further discussion of this issue). In addition, it is necessary to illustrate that each year (group) has the same parameter values if I wish to pool the samples (which I will do in the last section).

I can test the hypothesis that the parameters are the same in every year through a hierarchy of invariance, where parameters are constrained in an ordered sequence and compared to the unrestricted model. If the different years have similar parameter values, then there should be no significant difference in fit between the unrestricted model and the restricted models. In addition, the restricted models are nested within the unrestricted models, so a chi-square difference test provides a test of significance. Other fit statistics can also be checked for a serious decline in fit across the models.

For the trust component of the model, the parameters of interest are contained in the factor loadings, the variances and covariances of the latent variables, and the variances and covariances of the measurement errors. Therefore, I form the following hierarchy of models: the model with no restrictions on the values of the free parameters across time periods, the model with the factor loadings restricted across time, the model with both the factor loadings and the variance/covariance matrix of the latent variables restricted, and the model with all parameters restricted. The top half of table 2 indicates that there is no significant decline in fit when moving from the less restricted to the more restricted forms of the model.[33]

[32] I encountered some problems in the estimation of 1994. To achieve convergence, I had to constrain the variances of the measurement errors to those estimated in a multiple group analysis of the years 1975–93.

[33] The chi-square test statistics in the multiple group designs are all significant. This reflects the large sample sizes that result from combining all the years. The chi-square differences are not significant, however.

TABLE 2

HIERARCHY OF INVARIANCE FOR TWO COMPONENTS OF MODEL

| Model | $\chi^2$ | df | p-value | IFI | RMSEA | $\chi^2$ Difference | df | p-value |
|---|---|---|---|---|---|---|---|---|
| **Trust** | | | | | | | | |
| No constraints ...................... | 201.9 | 108 | 1E-07 | .08 | .025 | | | |
| Factor loadings constrained ... ........... ... ..... ... ..... | 252.1 | 148 | 2E-07 | .97 | .023 | 50.177 | 40 | .13 |
| Factor loadings, variance/covariance of latent vars constrained ... ....... . ...... . .... ... . ... .. .. ..... | 287.5 | 172 | 8E-08 | 97 | .022 | 35.361 | 24 | .063 |
| Factor loadings, var./cov. of latent vars, var./cov. of errors constrained . ..... . ......... .. ........ ..... | 317.2 | 236 | .0003 | .97 | .016 | 29.732 | 64 | 1 |
| **Associations·** | | | | | | | | |
| No constraints .......... ....... ...... ..... . ........ | 100 | 20 | 1.26E-12 | .97 | .050 | | | |
| Factor loadings constrained .. ... ..... .. ........ .. ..... .. | 111 | 26 | 1.78E-12 | .96 | .045 | 11 | 6 | .0884 |
| Factor loadings and structural paths constrained ....... .. | 133 | 50 | 1.83E-09 | .96 | .032 | 22 | 24 | .5793 |
| Factor loadings, structural paths, and var./cov. of error in equation constrained ... ..... ......... ...... ......... . | 138 | 58 | 1.81E-08 | .96 | .029 | 5 | 8 | .7576 |
| Factor loadings, structural paths, and all errors constrained ........ ... ...... ... ......... ..... ........ .. | 180 | 74 | 8.18E-11 | .95 | .030 | 42 | 16 | .0004 |

The chi-square difference test is not significant, and the other fit statistics show no decline in fit.[34]

In determining whether the parameters of the associations component remain stable, we face different parameters of interest. There, I consider the following hierarchy of invariance: the model with no restrictions on the values of the free parameters across time periods; the model with the factor loadings restricted across time; the model with both the factor loadings and the structural paths from the exogenous variables to the latent variable restricted; the model with the factor loadings, structural paths, and the variance of the error in the latent variable equation restricted; and the model with all the above and the variances and covariances of the measurement errors constrained across time. The bottom half of table 2 indicates that the parameters remain stable in the association component of the model. There is no significant decline in fit between any of the first four models. When the variances and covariances of the measurement errors are constrained across the years, there is a significant decline in the chi-square but the other fit indexes indicate no substantial decline in fit.

These results establish that there is no significant change in the parameters of the model over time. This means that it is safe to test for changes in the level of social capital over time, because the relationship between the general level of social capital and its indicators has remained stable over the 20-year period. The evidence for stable parameters also means that I can pool all the years together (one should not pool samples unless it is established that their parameters are the same), and use a pooled sample in my over-time tests.

Also contained within the hierarchy of invariance tables is the test for a change in dispersion. The fact that the variances of the latent variables have not changed over time indicates that there has been no change in the dispersion (variance) of social capital over time. Thus, although the mean level of social capital may show a decline over time, there is no evidence that it is becoming more unequal in its distribution. (An unequal distribution would mean that any change in the level of social capital was combined with increased inequality, indicating that some segments of the population are retaining or increasing their levels of social capital, while

[34] Again, because the endogenous variables were categorical, I analyzed the polychoric correlation matrix. As an additional check, I also estimated the trust component of the model with the AMOS maximum-likelihood missing value routine. There however, the chi-square statistic was significantly different between some of the models. The significance is almost certainly due to the power of the test, however, as the sample size of the multiple group test is over 12,000. This hypothesis was backed up in tests that showed the power to detect a number of standardized 0.1 changes to be 75%. The incremental fit index and the RMSEA, which are not affected by sample size, indicate no decline in fit.
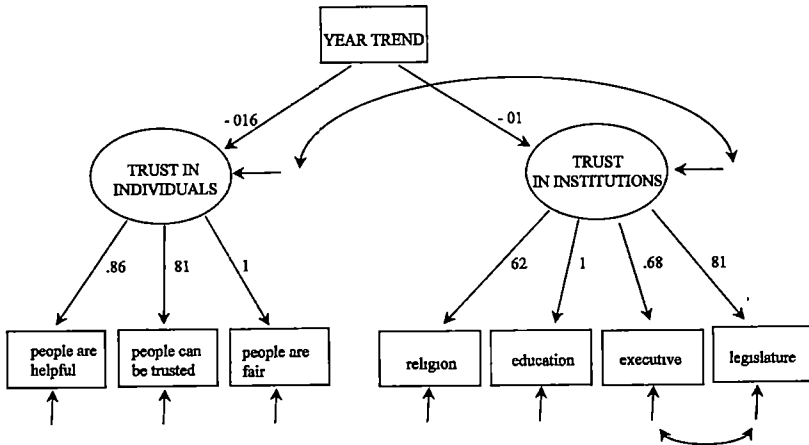
FIG. 6.—Model and estimated parameters for trust. All displayed coefficients are significant. $N = 9,148$; $\chi^2 = 343$; $df = 17$; $p$-value $= 0$; RMSEA $= .05$; IFI $= 1.0$.

others are decreasing in theirs. Since we do not see a change in the distribution of social capital over time, any decline in social capital we find is not accompanied by a concurrent increase in inequality.)

## Testing for a Decline in Social Capital

Having established that the model fits well in each year and that the estimated parameters remain stable over the time period, I can now estimate change in the level of social capital over time. The *GSS* samples a different set of individuals every year so I do not employ a traditional longitudinal design. Instead, I treat each year as a different "group" and consider change over time in social capital through a modification of a multiple group design. As outlined in Muthén (1989), differences in levels between groups can be estimated by pooling all of the groups and including exogenous dummy variables to distinguish between them. My ability to pool the years was established in the previous section. In this section, I pool all of the years and include a trend variable (YEAR TREND), which tracks the year of the observations beginning at 1 in 1975 and ending at 20 in 1994 (there are 19 years between 1975 and 1994).

Figures 6 and 7 present some of the relevant parameter estimates for the pooled sample with the yearly trend exogenous variable (YEAR TREND). Beginning with the trust component of the model: the negative, significant values from the trend variable to the two latent variables indi-
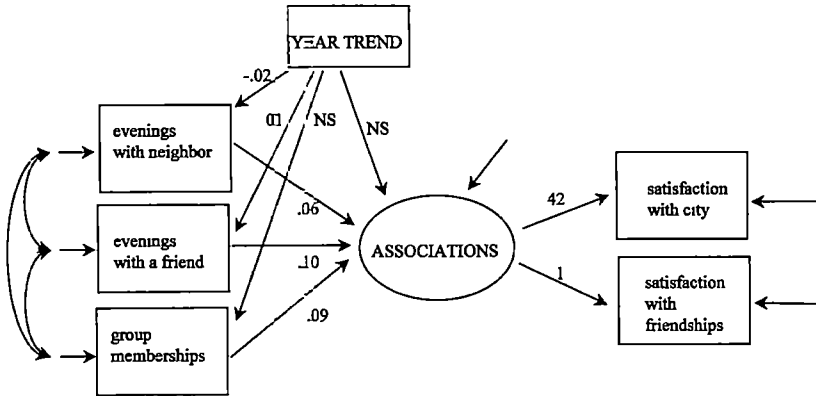
FIG. 7.—Estimated parameters of interest for associations. All displayed coefficients are significant. NS indicates a nonsignificant estimated parameter. $N =$ 14,268; $\chi^2 = 94$; $df = 3$; $p$-value $= 0$; RMSEA $= .05$; IFI $= .96$.

cate a negative trend over time. But we need to determine the actual effect size of these parameters. They are somewhat difficult to interpret because the coefficients come from a conditional polychoric analysis.[35] One way to interpret the effect size is to translate the coefficients into the change in the predicted probabilities of the observed variables (like in a probit analysis). These probabilities are provided as a chart over time in figures 8 and 9. For example, the change in the scaling indicator for trust in individuals over the 20-year period, when retransformed into percentages, is over 10% (from 0.49 to 0.37) or approximately a 0.5% drop per year. The indicators of trust in institutions, although also decreasing over time, are declining at a less rapid pace.

Things are quite different in the associations component of the model (figure 7). First, despite a very large pooled sample size, there is *no* significant change in the general level of association over time. I also assessed the impact of time on the three indicators that I expect to affect general levels of association. There is no change in individuals' memberships over time. There is a statistically significant 0.01 unit increase in the amount of time respondents claim they spend with friends outside their neighbor-

[35] For this component, since the trend variable is exogenous, I was able to compute the polychoric correlations conditional on YEAR TREND (Joreskog and Sörbom 1993, p. 180). Using that and the estimated variances of the variables provides LISREL with the covariance matrix for estimation rather than the polychoric correlation matrix. Thus, my parameter estimates are not standardized, as in a typical polychoric correlation analysis, but are unstandardized estimates of effects.
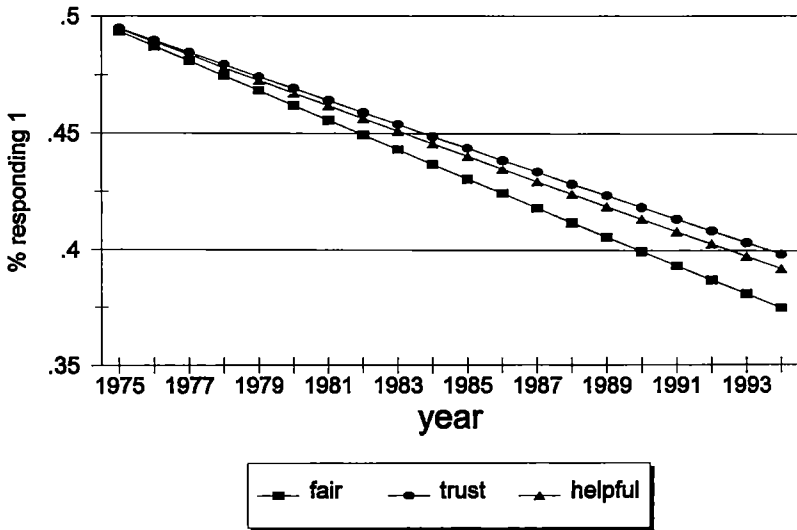
FIG. 8.—Predicted probabilities: trust in individuals
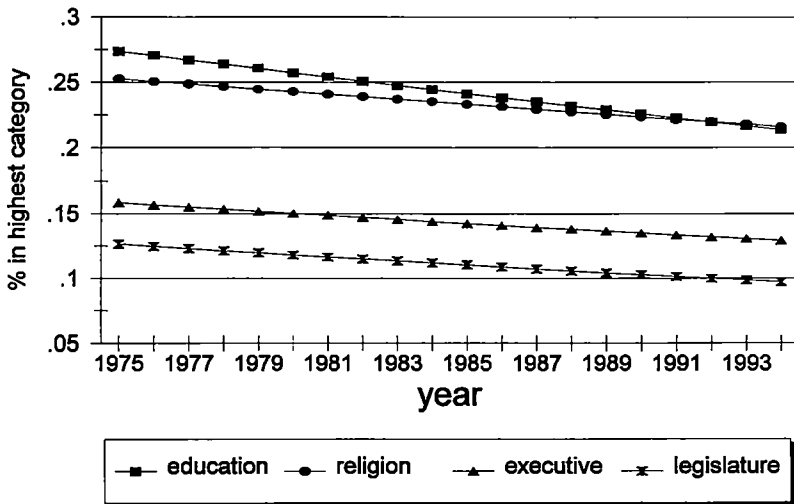


FIG. 9.—Predicted probabilities: trust in institutions
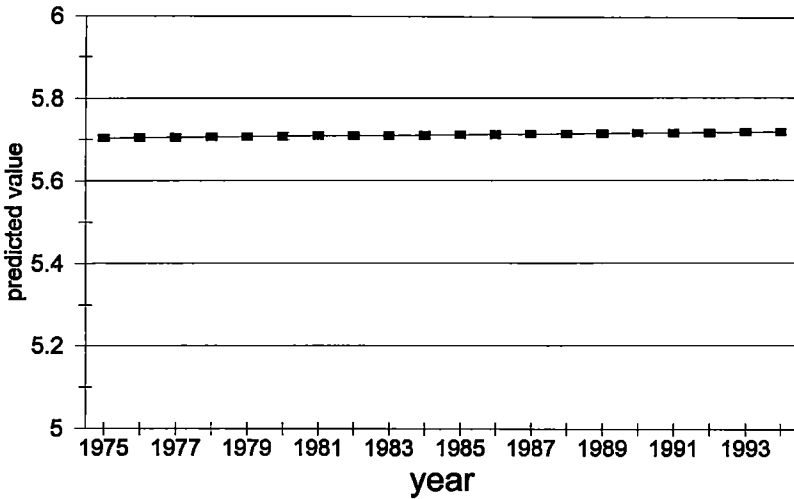
FIG. 10.—Predicted value: latent associations

hood and a 0.02 unit decrease in the time they claim to spend with neighbors. The actual effect size is extremely small, however. The range of the two variables is from one to seven, so the decline, while statistically significant, is not practically significant. To further illustrate the lack of change over time, in figure 10, I provide the predicted value of the unobserved latent variable as a chart over time. The chart indicates that, clearly, the latent level of association has not changed over time.

To briefly review, the analysis of change in social capital has shown a decline in trust and no change in associations. There are two other issues that must be considered, however, before moving to an analysis of general social capital. First, measuring time as a single trend variable assumes that time has a linear effect on the latent variables. To relax that assumption, I can utilize a series of dummy variables to represent individual years. If the dummy variables are coded to represent their year in the trend rather than the typical 0/1 coding (e.g., 0/0, 0/3, 0/8 rather than 0/0, 0/1, 0/1), then they can be used to test for nonlinear trends.[36] A significant chi-square difference between the restricted and unrestricted models means that a nonlinear trend is present.

---

[36] A restricted version of the model with multiple dummy variables, where all the estimated coefficients for the dummy variables are restricted to be the same, will produce the same results as a model with a single trend variable. When the constraints on the estimation of the coefficients are relaxed, however, the model allows nonlinear effects.
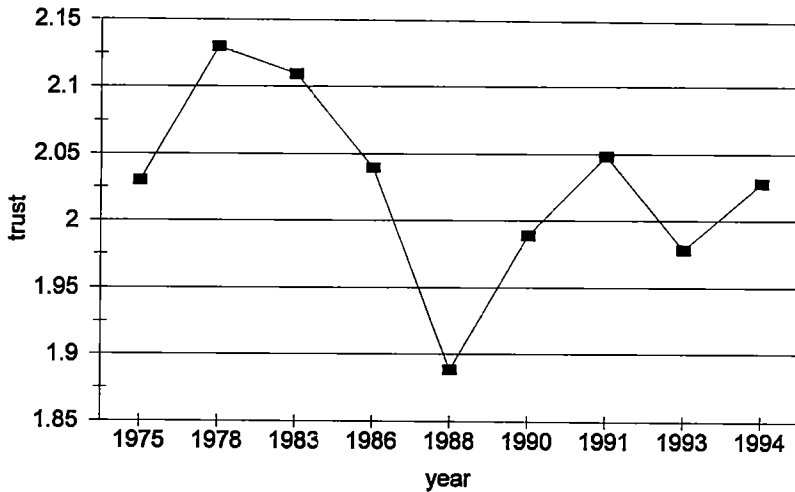
FIG. 11.—Trust in religion: 1975–94

Both components of social capital, trust and associations, do show evidence of a nonlinear trend. The chi-square difference test between the model with constrained dummy variable coefficients and free coefficients is 250 with 14 degrees of freedom for the trust component of the model. The chi-square difference test for the associations component is 49 with 28 degrees of freedom. Both tests indicate a significant improvement in fit when the dummy variables are allowed to estimate a nonlinear trend.[37]

A second issue that needs to be addressed is whether trust in specific institutions could be affected by yearly events, as well as over-time changes in general trust.[38] That is, scandals in a particular year, related to a particular institution, could drive responses about trust in that institution in that year. For example, consider the chart of the mean of trust in organized religion over time (fig. 11). The trend is dominated by the decline in trust in 1988, which is the year immediately following the Jim

[37] The trends for trust in individuals and institutions are discussed in more detail below The trends for association memberships indicate that the general level of association does not really change over the time period, and neither does number of group memberships. The other two indicators show opposite trends. Neither shows any decline in the earliest years, but after 1988, spending an evening with a community friend steadily declines. Spending an evening with a noncommunity friend actually increases in the last few years.

[38] This is comparable to Smith's (1997) research, which illustrates that individual life events, such as criminal victimization, can influence people's responses to trust in individuals.

Baker (television evangelist) and related religious scandals. Another example of a potentially important scandal is Watergate, which could influence individuals' responses about trust in government institutions.[39] To assess these types of effects, I can use multiple dummy variables and allow the intercepts of indicators to be influenced by particular years (Muthén 1989). So, I can model a specific yearly effect by including a path from the dummy variable for a specific year directly to trust in a particular institution.

The first step in such an analysis is to identify scandals that are potentially relevant to the four institutions in my model. To determine significant scandal events surrounding the institutions of my model, I considered the *World Almanac and Book of Facts* (1975–94) list of top news stories for each year. The *GSS* is conducted in February, March, and April of every year, so I used the top news stories for the previous year in determining the potential exogenous effect of a particular year on any given institution. The *World Almanac* suggested four paths. The first, from 1988 to trust in religious organizations, reflects the Jim Baker and related religious scandals of 1987. Another path, from 1988 to the executive branch of the government, reflects the Iran-Contra affair. Finally, two paths, from 1978 to the legislature and the executive, reflect possible lingering effects of the Watergate scandal (1975 was not used because it is the omitted dummy category). Figure 12 graphically represents how I test for a nonlinear trend and the effects of scandals in the trust component of the model. I include a series of dummy variables for the test of nonlinear effects (eight variables representing the nine years for which I have data, minus the omitted year, 1975) and the specific yearly effects included to account for scandals (from a specific year, say 1988, to a specific institution, say religion).

Of the specific yearly events included in the model, all are significant except the path from 1978 to trust in the legislature. More important, once the paths related to scandals in particular years are included, there is no longer a consistent downward trend in the general (latent) level of trust in institutions over the time period. Instead, there is first increasing trust in institutions over the time period and then decreasing trust.[40] Therefore, when shocks to trust in institutions related to specific events are allowed

---

[39] For more information, see Smith (1994), who outlines the distinction between slow, over-time opinion change and event-driven change. One of his examples is the televangelist scandal of 1988. Most change is slow and consistent—it is this opinion change that we want to distinguish from sharp, event-driven changes.

[40] Without the scandal years in the model, trust in individuals has a general downward trend, while trust in institutions shows little change across most of the time period, with declines toward the end. After scandal years are included, general trust in institutions actually increases slightly over the time period, with large decreases in the later years
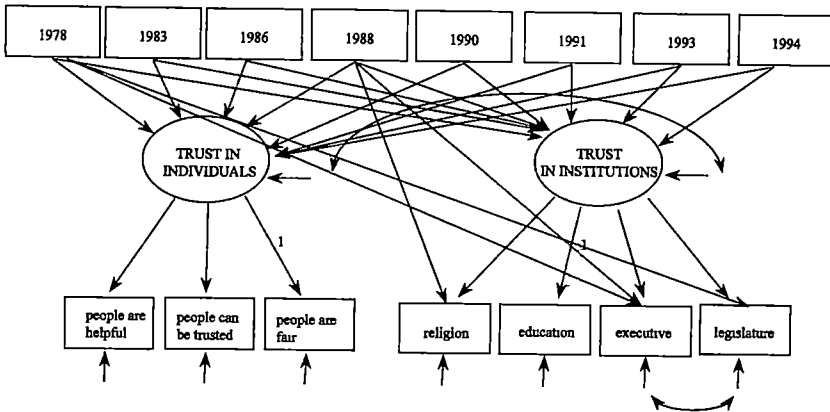
FIG. 12.—Model to test for nonlinear and scandal effects

in the model, there remains no separate general decline in trust in institutions.[41]

The above analyses consider the two components of social capital separately. The definition of social capital I presented earlier, however, describes social capital as the *combination* of trust and associations. I must therefore also track social capital, as a combination of both components, and check for a decline over time. There are a number of strategies for combining trust and associations, so to be more comprehensive, I present two alternative measures. First, I multiply the means of the relevant latent variables, providing a picture of their combined movement over time. Second, I track the percentage of individuals who both trust and associate at high levels.

For the trust component of the model, determining the mean of a latent variable (either trust in individuals or trust in institutions) in any year, $\kappa_{it}$, is straightforward—it is the mean of its scaling indicator for that year. However, because the associations model is not a traditional confirmatory factor analysis, the mean of latent associations is more complicated to calculate. There, in each year, the formula is:

$$E(\eta) = (\alpha + \Gamma\mu_x),$$

[41] The asymptotic covariance (weight) matrix for the trust component of the model with multiple dummy variables was nonpositive definite. Therefore, that portion of the analysis could not be estimated in LISREL with the conditional correlation matrix, as before. Instead, I estimated it in AMOS under the assumption that the indicators are continuous.

119
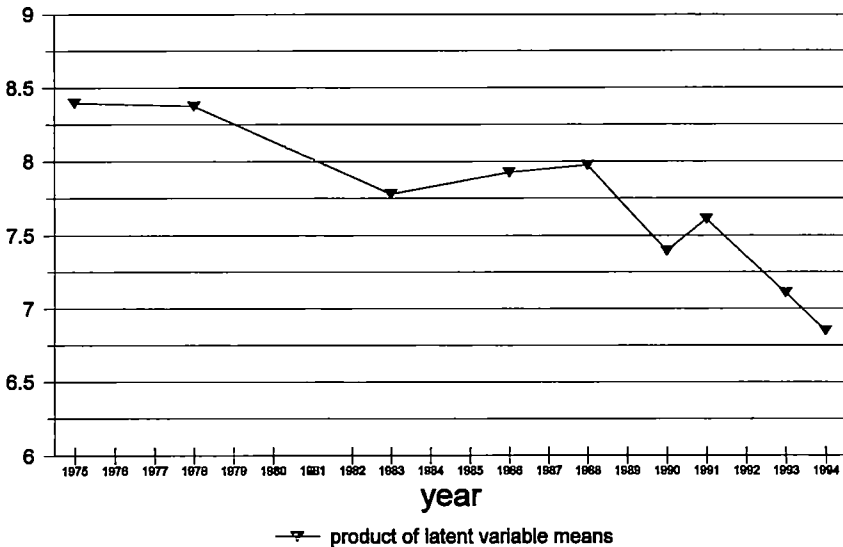
FIG. 13.—Latent social capital over time

where $E(\eta)$ is the mean of the latent association variable, $\alpha$ is the intercept of the latent variable, $\Gamma$ is the vector of coefficients representing the impact of the exogenous x's on the latent variable, and $\mu_\chi$ is the means of the exogenous x's. For each year, I estimated the means of the latent variables (latent trust in individuals, latent trust in institutions, and latent associations), and the line in figure 13 presents their product, tracked over time. It shows a slight decline over the time period, from 8.4 to 6.9.

As an additional measure, at the individual level, we can determine what percentage of the population displays both trust and associations over time. I selected individuals who responded "yes" to all three questions about trust, had at least "some" confidence in all four institutions, and belonged to at least one group. The line in figure 14 displays these percentages over time. It shows a sharper decline in social capital over time, from about 15% to 9%. As illustrated by the analyses of the two components separately, the downward trend in social capital is closely related to declines in the trust component. Over time, neither the mean of latent associations nor the percentage of individuals belonging to at least one association shows a decline. As the combination of both trust and associations, however, social capital shows declines over the 20-year period.

year

—▼— % trusting and associating indivs.
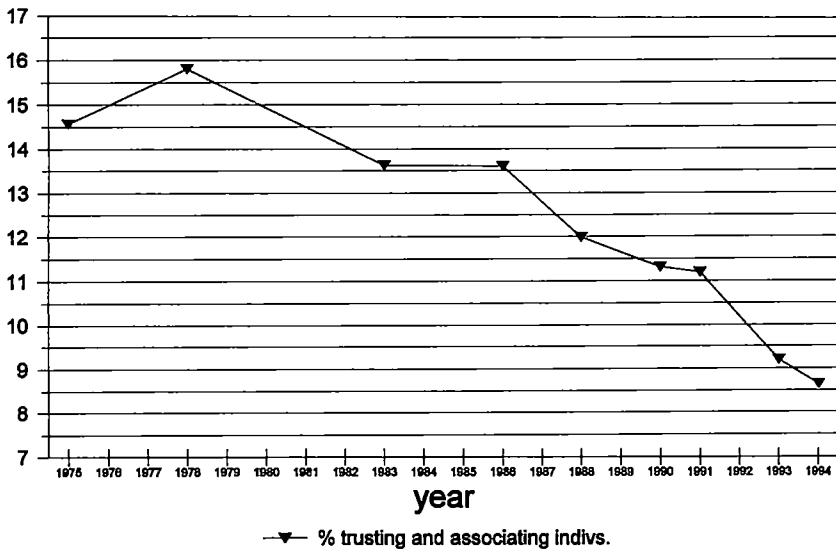
Fig. 14.—Individual social capital over time

CONCLUSION

This article has attempted to expand upon both theoretical discussions of social capital and empirical analyses of its potential decline in the United States. In the article, I present an empirical model of social capital that corresponds to two theoretically outlined components—trust and associations. I also distinguish between social capital at the national or community level and social capital at the individual level and within-group level. My empirical model contains multiple indicators of each component and accounts for error in the measurement of these indicators. I test for a linear decline in social capital over a 20-year period. In addition, I estimate a more flexible model that uses dummy variables for each year and considers specific yearly effects on some indicators.

In summary, my results do not consistently support Putnam's claim of a decline in social capital. I do find that my measure of social capital, as a combination of trust and associations, shows a decline over the time period. This is mainly due to a strong and consistent decline in trust in individuals over the period 1975–94—about a 0.5% drop per year. I do not find a general decline in trust in institutions, however, once scandals in particular years are included. While discussions of trust in institutions anecdotally mention the importance of specific events (e.g., Lipset and Schneider 1983, p. 399; Wuthnow 1997, p. 15), none empirically estimate

their effects. The difference between trust in specific institutions and a more general level of trust in institutions is important to distinguish in this literature. The theory of social capital is concerned with trust at the more general level (closer to legitimacy), and so it is important to note that this general trust in institutions has not declined over the period even though specific institutions were influenced by scandals.

While trust in individuals has declined over the time period, the second component of social capital, the level of associations, remains unchanged. Membership in groups has not declined, and there has been little practical change in the amount of time that individuals spend with neighbors and friends outside the neighborhood. There is some evidence for a shift toward more associations outside of neighborhoods, which is corroborated by Wuthnow (1997, p. 41) who shows that more people are looking outside of their neighborhoods for new friends. Whether this slight trend toward associating outside the neighborhood will decrease social capital by decreasing the quality of interaction or increase it by expanding crosscutting ties requires further research.

This article provides two additional results. First, my analysis of the variance of social capital shows no general change in the dispersion of social capital over time. However, Wuthnow (1997) makes the important observation that a decline in the associational component of social capital is not evenly spread among the population but is instead concentrated in "marginalized" sections of the population. Assessing differences in the amount of social capital held by different segments of the population and how that distribution may have changed over time is essential future research.

Second, this article provides evidence that the relationship between the indicators of social capital and the theoretical concept of social capital has not changed over time. This information should make us feel safer when assessing social capital in time periods for which we do not have multiple indicators. For example, some have argued that general declines in trust began in the 1960s. To test such an assumption, we must utilize a longer time series. The questions on trust in individuals have been asked in various surveys since 1964. Since the interpretation of those questions was stable in the time period I consider, 1975–94, it is likely that it is stable in time periods before 1975 as well. Trends in those questions before 1975 do indicate that the decline in trust began even before the time period I consider (see Smith [1997] for a compilation of data for the longer time period).[42]

---

[42] While Smith (1997) is directly interested in a measure of "misanthropy" over time, his article provides data on trust over a long period of time. The data display a long-

In considering the implications of these results, we must remember that this article was not a simple test of Putnam's thesis. Instead, the question of a decline in social capital ties into classic concerns in sociology about declines in community and their potentially detrimental consequences. If our theories are correct, then a decline in social capital could mean problems for the maintenance of U.S. democracy in the future.[43]

It appears as yet that we do not have to worry about a decline in associations or a general decline of trust in institutions. However, the strong, consistent decline in trust in individuals (and its effect on social capital as the combination of trust and associations) could have some potentially detrimental consequences. For example, consider the recent rise of gated communities and the increased use of private security guards. It could be that our trust in one another impacts how we organize our lives and how we choose to spend our money (or how much money we spend). Or, consider the recent article in *Science* (Sampson, Raudenbush, and Earls 1997) that showed that trust, as one component of collective efficacy, is linked to reduced violence in neighborhoods. As we consider the *effects* of social capital in more areas, we can better understand the consequences of a *decline* in social capital. For example, Wilson and Musick (1997) found that social capital was related to higher rates of volunteering. Finally, there is always the possible negative impact on democracy from a decline in trust.

While this study represents an improved estimate of national-level social capital, researchers need to construct and obtain more precise measures of social capital for present and future assessments. Most important, a measure of the crosscutting ties between associations is sorely needed. Without such a measure, we can not fully establish whether U.S. involvement in associations produces *positive* social capital for the nation. Also, this article focused on the measurement of social capital and changes in its level over time. It did not consider social capital in subgroups of the population or *reasons* for changes in social capital over time. Further research is therefore needed to understand the determinants of social capital (e.g., Brehm and Rahn 1997; Wuthnow 1997) and whether it can be transferred or "infused" from areas or groups with high social capital to groups with low social capital.

The idea of social capital is linked to many classical theories of what

---

term decline in the variables "trust" and "fair." "Helpful" shows a more variable pattern in both the *GSS* data and non-NORC sources over time.

[43] One important question is whether a certain "threshold" of social capital is necessary for the successful maintenance of democracy. It may be that public goods remain undisturbed until social capital falls below some level. Determining the presence and value of such thresholds would be interesting future research.

makes a "good" society. Understanding this, we should continue to monitor the level of social capital in the United States. While this article has not shown a consistent decline in the level of both components of social capital over the last 20 years, their combination *has* declined. And, the possibility of change in the future should not be ignored. We should also consider ways to protect and enrich our stock of social capital to ensure a healthy society in the future.

REFERENCES

Arbuckle, James L. 1995. AMOS, ver. 3.5. Chicago: SmallWaters Corporation.
Arendt, Hannah. 1948. *The Origins of Totalitarianism.* New York: Harcourt & Brace.
Argyle, Michael. 1992. "Benefits Produced by Supportive Social Relationships." Pp. 13–31 in *The Meaning and Measurement of Social Support,* edited by Hans O. F. Veiel and Urs Baumann. New York: Hemisphere.
Banfield, Edward C. 1958. *The Moral Basis of a Backward Society.* Glencoe, Ill.. Free Press.
Barber, Bernard. 1983. *The Logic and Limits of Trust.* New Brunswick, N.J.: Rutgers University Press.
Baumgartner, Frank R., and Jack L. Walker. 1988. "Survey Research and Membership in Voluntary Associations." *American Journal of Political Science* 32:908–27
Becker, Gary Stanley. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* New York: National Bureau of Economic Research.
———. 1996. *Accounting for Tastes.* Cambridge, Mass.: Harvard University Press.
Bollen, Kenneth A. 1989a. *Structural Equations with Latent Variables.* New York: Wiley.
———. 1989b. "A New Incremental Fix Index for General Structural Equation Models." *Sociological Methods and Research* 17:303–16.
Bourdieu, Pierre. 1983. "Forms of Capital." Pp. 241–58 in *Handbook of Theory and Research for the Sociology of Education,* edited by John G. Richardson. New York: Greenwood Press
Bourdieu, Pierre, and Loïc Wacquant. 1992. *An Invitation to Reflexive Sociology.* Chicago: University of Chicago Press.
Brehm, John, and Wendy Rahn. 1997. "Individual-Level Evidence for the Causes and Consequences of Social Capital." *American Journal of Political Science* 41:99–1023.
Breiger, Ronald. 1974. "The Duality of Persons and Groups." *Social Forces* 53:181–90.
Browne, Michael, and Robert Cudeck. 1993. "Alternative Ways of Assessing Model Fit." Pp. 136–62 in *Testing Structural Equation Models,* edited by Kenneth A Bollen and J. Scott Long. Newbury Park, Calif.: Sage.
Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94 S95–S120.
———. 1990. *Foundations of Social Theory.* Cambridge, Mass.. Harvard University Press.
Dasgupta, Partha. 1988. "Trust as a Commodity." Pp. 49–72 in *Trust: Making and Breaking Cooperative Relationships,* edited by Diego Gambetta New York: Basic Blackwell.

Davis, James Allen, and Tom W. Smith. 1994. *General Social Surveys, 1972–1994* (MRDF). Chicago: National Opinion Research Center.

de Tocqueville, Alexis. [1835, 1840] 1990. *Democracy in America.* New York: Vintage.

Ehrenholt, Alan. 1995. "No Conservatives Need Apply." *New York Times,* November 19.

Fatton, Robert Jr. 1995. "Africa in the Age of Democratization: The Civic Limitations of Civil Society." *African Studies Review* 38:67–99.

Foley, Michael W., and Bob Edwards. 1997. "Escape from Politics: Social Theory and the Social Capital Debate." *American Behavioral Scientist* 40:550–61.

Fukuyama, Francis. 1995 *Trust: The Social Virtues and the Creation of Prosperity.* New York: Free Press.

Gambetta, Diego. 1988. "Can We Trust Trust?" Pp. 213–37 in *Trust: Making and Breaking Cooperative Relationships,* edited by Diego Gambetta. New York: Basic Blackwell.

Giddens, Anthony. 1990. *The Consequences of Modernity.* Stanford, Calif.: Stanford University Press.

Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360–80.

———. 1985. "Economic Action and Social Structure  The Problem of Embeddedness." *American Journal of Sociology* 91:481–510.

Guest, Avery M., and R S. Orpesa. 1986. "Informal Social Ties and Political Activity in the Metropolis." *Urban Affairs Quarterly* 21:550–74.

Hagan, John, Ross MacMillan, and Blair Wheaton. 1996. "New Kid in Town: Social Capital and the Life Course Effects of Family Migration on Children." *American Sociological Review* 61:368–85.

Halebsky, Sandor. 1976. *Mass Society and Political Conflict: Toward a Reconstruction of Theory.* Cambridge: Cambridge University Press.

Hanifan, L. J. 1920. *The Community Center.* Boston: Silver, Burdette & Co.

Hardwig, John. 1991. "The Role of Trust in Knowledge." *Journal of Philosophy* 88: 693–708.

Jacobs, Jane. 1961. *The Death and Life of Great American Cities.* New York· Random House.

Jöreskog, K. G., and D. Sörbom. 1986. *LISREL 6: Analysis of Linear Structural Relationships by Maximum Likelihood and Least Square Methods.* Mooresville, Ind.: Scientific Software.

———. 1993. *LISREL 8: User's Reference Guide.* Chicago: Scientific Software International.

Ladd, Everett C. 1996. "The Data Just Don't Show Erosion of America's Social Capital." *Public Perspective* 7:1–30.

Lee, Barrett A., R. S. Oropesa, Barbara J. Metch, and Avery M. Guest. 1984. "Testing the Decline-of-Community Thesis: Neighborhood Organizations in Seattle, 1929–1979." *American Journal of Sociology* 89:1161–88.

Lehman, Edward. 1987. "The Crisis of Political Legitimacy: What Is It; Who's Got It; Who Needs It." *Research in Political Sociology* 3:203–21.

Lemann, Nicholas. 1996. "Kicking in Groups." *Atlantic Monthly* 277 (4): 22–26.

Lipset, Seymour Martin, and William Schneider. 1983. *The Confidence Gap: Business, Labor, and Government in the Public Mind.* New York: Free Press.

Lipset, Seymour Martin, Martin A. Trow, and James S. Coleman. 1956. *Union Democracy: The Internal Politics of the International Typographical Union.* Glencoe, Ill.: Free Press.

Loury, Glenn. 1977. "A Dynamic Theory of Racial Income Differences." Pp. 153–88 of *Women, Minorities, and Employment Discrimination,* edited by P. A. Wallace and A. LeMund. Lexington, Mass.: Lexington Books.

Luhmann, Niklas. 1979. *Trust and Power.* Chichester: Wiley.

―――. 1988 "Familiarity, Confidence, Trust: Problems and Alternatives." Pp 94–107 in *Trust: Making and Breaking Cooperative Relationships,* edited by Diego Gambetta. New York: Basic Blackwell.

McPherson, J. Miller. 1982 "Hypernetwork Sampling: Duality and Differentiation among Voluntary Associations." *Social Networks* 3·225–49.

Mill, John Stuart. (1831) 1975. *Three Essays.* Oxford: Oxford University Press.

Moody, James, and Peter S. Bearman. 1997. "Shaping School Climate: School Context, Adolescent Social Networks, and Attachment to Schools." Manuscript. University of North Carolina at Chapel Hill, Department of Sociology.

Muthén, Bengt O. 1989. "Latent Variable Modeling in Heterogeneous Populations." *Psychometrika* 54:557–85.

Nelson, Joan M. 1994. "Labor and Business Roles in Dual Transitions: Building Blocks or Stumbling Blocks?" Pp. 147–94 in *Intricate Links: Democratization and Market Reforms in Latin America and Eastern Europe,* edited by Joan M. Nelson. New Brunswick, N.J.: Transaction Publishers.

Paxton, Pamela. 1998. "Capitalizing on Community· Social Capital and the Democratic Society." Ph.D. dissertation. University of North Carolina at Chapel Hill, Department of Sociology

Portes, Alejandro. 1998. "Social Capital: Its Origins and Applications in Modern Sociology." *Annual Review of Sociology* 22:1–24.

Portes, Alejandro, and Julia Sensenbrenner. 1993. "Embeddedness and Immigration: Notes on the Social Determinants of Economic Action." *American Journal of Sociology* 98:1320–50.

Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy.* Princeton, N.J.: Princeton University Press.

―――. 1995. "Bowling Alone: America's Declining Social Capital." *Journal of Democracy* 6:65–78.

―――― 1996. "The Strange Disappearance of Civic America." *American Prospect* 24: 34–48.

Reiman, J. 1990. *The Rich Get Richer and the Poor Get Prison,* 3d ed. New York: Macmillan.

Rosenberg, Morris. 1956. "Misanthropy and Political Ideology." *American Sociological Review* 21:690–95.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277: 918–24

Sanders, Jimy M., and Victor Nee. 1996. "Immigrant Self-Employment: The Family as Social Capital and the Value of Human Capital." *American Sociological Review* 61:231–49.

Schiff, Maurice. 1992. "Social Capital, Labor Mobility, and Welfare." *Rationality and Society* 4:157–75.

Schultz, Theodore W. 1961. "Investment in Human Capital." *American Economic Review* 51:1–17.

Simmel, Georg. (1903) 1950. "The Metropolis and Mental Life." Pp. 409–24 in *The Sociology of Georg Simmel,* translated and edited by Kurt H Wolff. New York: Free Press.

―――. 1971. *On Individuality and Social Forms.* Chicago: University of Chicago Press.

Smith, Mark H., Lionel J. Beaulieu, and Ann Seraphine. 1995. "Social Capital, Place of Residence, and College Attendance." *Rural Sociology* 60:363–80.

Smith, Tom W. 1981. "Can We Have Confidence in Confidence: Revisited." Pp. 119–89 in *Measurement of Subjective Phenomena,* edited by Denis F. Johnston. Washington, D.C · U.S. Department of Commerce.

———. 1990. "Trends in Voluntary Group Membership: Comments on Baumgartner and Walker." *American Journal of Political Science* 34:646–61.

———. 1994. "Is There Real Opinion Change?" *International Journal of Public Opinion Research* 6:187–203

———. 1997. "Factors Relating to Misanthropy in Contemporary American Society." *Social Science Research* 26:170–96.

Steiger, J. H., and J. M. Lind. 1980. "Statistically Based Tests for the Number of Common Factors." Paper presented at the annual meeting of the Psychometric Society, Iowa City, Iowa.

Stein, Maurice. 1960 *The Eclipse of Community: An Interpretation of American Studies.* New York: Harper & Row.

Tanaka, J. S. 1993. "Multifaceted Conceptions of Fit in Structural Equation Models." Pp. 10–39 in *Testing Structural Equation Models,* edited by Kenneth A. Bollen and J. Scott Long. Newbury Park, Calif.: Sage Publications.

Tönnies, F. (1887) 1957. *Community and Society—Gemeinschaft und Gesellschaft,* translated and supplemented by C. P. Loomis. East Lansing: Michigan State University Press.

Uzzi, Brian. 1996. "The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect." *American Sociological Review* 61:674–98.

Verba, Sidney, Norman H. Nie, and Jae-On Kim. 1978 *Participation and Political Equality: A Seven-nation Comparison.* Cambridge: Cambridge University Press.

Verba, Sidney, Kay Lehman Schlozman, and Henry E. Brady. 1995. *Voice and Equality: Civic Volunteerism in American Politics.* Cambridge, Mass.: Harvard University Press.

Wasserman, Stanley, and Katherine Faust. 1994 *Social Network Analysis: Methods and Applications.* Cambridge: Cambridge University Press.

Wellman, Barry, Peter J. Carrington, and Alan Hall. 1988. "Networks as Personal Communities." Pp. 130–84 in *Social Structures: A Network Approach,* edited by Barry Wellman and S D. Berkowitz. Cambridge: Cambridge University Press.

Wilson, John, and Marc Musick. 1997. "Who Cares? Toward an Integrated Theory of Volunteer Work." *American Sociological Review* 62:694–713.

Wirth, Louis. 1938. "Urbanism as a Way of Life " Pp. 60–83 in *On Cities and Social Life,* edited by Albert J Reiss, Jr. Chicago: University of Chicago Press.

Wolfinger, Raymond, and Steven Rosenstone. 1980. *Who Votes?* New Haven, Conn.: Yale University Press.

Woolcock, Michael. 1998. "Social Capital and Economic Development: Towards a Theoretical Synthesis and Policy Framework." *Theory and Society* 27:151–208.

*World Almanac and Book of Facts.* 1975–94. New York: Newspaper Enterprise Association.

Wuthnow, Robert. 1997. "The Changing Character of Social Capital in the United States." Working paper. Princeton University, Department of Sociology.

Zhou, Min, and Carl Bankstron III. 1994. "Social Capital and the Adaptation of the Second Generation: The Case of Vietnamese Youth in New Orleans." *International Migration Review* 28:841–45.

# Structure and Strategy in Collective Action[1]

Michael Suk-Young Chwe
*University of Chicago*

This article considers both structural and strategic influences on collective action. Each person in a group wants to participate only if the total number taking part is at least her threshold; people use a network to communicate their thresholds. People are strategically rational in that they are completely rational and also take into account that others are completely rational. The model shows first that network position is much more important in influencing the revolt of people with low thresholds than people with high thresholds. Second, it shows that strong links are better for revolt when thresholds are low, and weak links are better when thresholds are high. Finally, the model generalizes the threshold models of Schelling (1978) and Granovetter (1978) and shows that their findings that revolt is very sensitive to the thresholds of people "early" in the process depends heavily on the assumption that communication is never reciprocal.

## INTRODUCTION

Collective action has been studied in two largely disjoint approaches, one focusing on the influence of social structure and another focusing on the incentives for individual participation. These approaches are often seen

as competing or even opposed. This article describes a simple model that tries to bridge this methodological division, using concepts from both social network theory and game theory. Here, a group of people face a collective action problem in that an individual wants to participate only if joined by enough others; exactly how many total participants are necessary is given by the individual's "threshold." Individuals are located in a social network, and each person knows the thresholds of only her neighbors in the network; each person has "local knowledge." People are strategically rational: they are completely rational and make decisions knowing that others are completely rational.

The model makes three substantive points. First, many empirical studies of collective action make the assumption, implicit in linear regression, that a person's individual characteristics and social position enter into participation linearly and independently. My model suggests that this might not be valid. In this model, people with low thresholds, who are highly predisposed toward participation, are affected much more by social position than people with high thresholds. Intuitively, whether a low-threshold person participates or not depends greatly on whether that person happens to have some sympathetic friends, while a high-threshold person participates only if a great mass of people participate. Second, although widely scattering "weak links" seem to be better for widespread communication than more involuted "strong links" (Granovetter 1973), empirical researchers have found that strong links, not weak links, correlate positively with participation (e.g., McAdam 1986; McAdam and Paulsen 1993; Valente 1995). Our model helps resolve this puzzle by showing that when thresholds are low, strong links can be better for participation. Weak links are better at spreading information widely, but strong links are better at locally creating the common knowledge, that is, knowledge of other people's knowledge, essential for collective action. Third, the model generalizes the threshold models of Schelling (1978) and Granovetter (1978) and shows that their finding that collective action is very sensitive to the thresholds of people "early" in the process depends heavily on the assumption that communication is never reciprocal. When even a small possibility of reciprocal communication is allowed, collective action is fairly robust. These three points are made using randomly generated networks of 30 people. I conclude by discussing the model in relation to existing models, especially on the issue of network transitivity.

## THE MODEL

There is a group of $n$ people, and each person chooses either to revolt $r$ (participate in the collective action) or stay at home $s$ (not participate). Each person $i$ has an idiosyncratic threshold $\theta_i \in \{1, 2, \ldots, n + 1\}$; a person wants to revolt only if the total number of people who revolt is greater than or equal to her threshold. For example, a person with threshold 2 prefers to revolt if he is joined by at least one other; a person with threshold $n$ prefers to revolt only if everyone else does. The social network $\rightarrow$ is a binary relation over $N$, where $j \rightarrow i$ means that person $j$ talks to person $i$. We define $B(i) = \{j \in N: j \rightarrow i\}$ to be person $i$'s "neighborhood," the set of people who talk to $i$. We assume that $\rightarrow$ is reflexive $(i \rightarrow i)$ and thus $i \in B(i)$. The idea is that person $i$ knows the thresholds of only the people in her neighborhood $B(i)$. We also assume that person $i$ knows all network relations among the people in $B(i)$; in other words, for all $j$, $k \in B(i)$, he knows whether or not $j \rightarrow k$.

To model this as a game, we have to specify information, strategies, and payoffs and define equilibrium. The details of this are in the appendix; here, I try to explain the game using some simple examples. The key modeling principle here is that a person's knowledge determines his ability to distinguish between states of the world, and if a person cannot distinguish between several states of the world, he must take the same action in all of them. For example, say there are only two people: person 1 has a threshold of either 1, 2, or 3, and person 2 also has a threshold of either 1, 2, or 3. Hence, there are nine possible states of the world: 11, 12, 13, 21, 22, 23, 31, 32, 33, where 23 is the state in which person 1 has threshold 2 and person 2 has threshold 3, for example.

Say that person 1 and person 2 do not communicate; we have the "null network," as shown in figure 1 (of course $1 \rightarrow 1$ and $2 \rightarrow 2$, but throughout we leave out these "loops" for clarity). Hence, each person only knows his own threshold. We can represent person 1's knowledge by the sets {11, 12, 13}, {21, 22, 23}, {31, 32, 33}, which form a partition of the nine possible states of the world, as shown in figure 1. The idea here is that if two states of the world are in the same set, then she cannot distinguish between them; for example, person 1 cannot distinguish between states 21, 22, and 23 because she does not know person 2's threshold. Similarly, since person 2 only knows his own threshold, his partition is {11, 21, 31}, {12, 22, 32}, {13, 23, 33}, also shown in figure 1.

Person 1 chooses whether to revolt or stay at home given a state of the world. If her threshold is 1, then she is happy to revolt all by herself and thus chooses to revolt in states 11, 12, and 13, as shown in figure 1. If her threshold is 3, then she never wants to revolt under any circumstances and thus chooses to stay at home in states 31, 32, and 33. If her threshold

Person 1 •          • Person 2

| 11 | 12 | 13 |
|----|----|----|
| 21 | 22 | 23 |
| 31 | 32 | 33 |

Person 1's partition

| 11 | 12 | 13 |
|----|----|----|
| 21 | 22 | 23 |
| 31 | 32 | 33 |

Person 2's partition

| r | r | r |
|---|---|---|
| s | s | s |
| s | s | s |

Person 1's actions

| r | s | s |
|---|---|---|
| r | s | s |
| r | s | s |

Person 2's actions
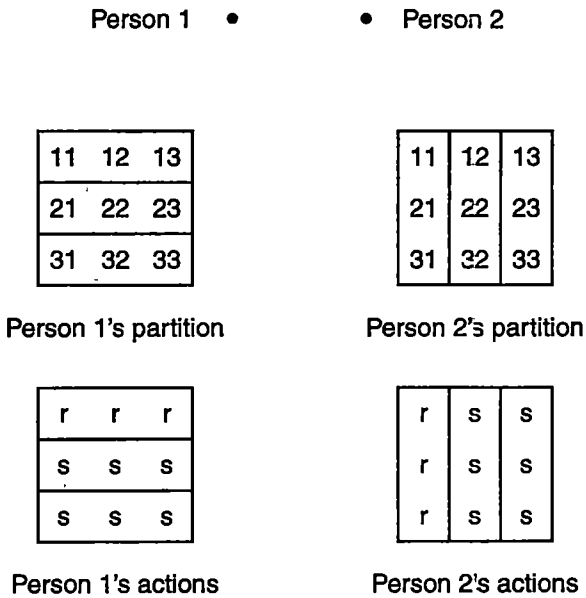
FIG. 1.—Null network (no communication)

is 2, then ideally person 1 would like to revolt in state 21 (in which case person 2 will surely revolt) and stay at home in state 23 (in which case person 2 will surely stay at home). But she cannot distinguish between these states; she does not know person 2's threshold. If she cannot distinguish between two states, then she cannot condition her action on them. In other words, she must revolt in all three states 21, 22, and 23 or stay at home in all three states. If she chooses to revolt, then person 2 might revolt (if the state is 21) but might not (if the state is 23). In other words, if she revolts, then there is the possibility that the total number of people revolting is less than her threshold. We assume (for the sake of modeling simplicity) that a person gets a very large negative payoff or penalty if this happens and hence a person revolts only if she knows for certain that enough others will revolt. Hence, person 2 decides to stay at home in states 21, 22, and 23, as shown in figure 1. Similarly, person 2's actions are also shown in figure 1.

Now consider the case when we have the "complete graph" network in which 1 → 2 and 2 → 1, as shown in figure 2 (when an arc is symmetric, we leave out the arrows for convenience). Here, each person knows each other's threshold. Now person 1's partition (and also person 2's) is {11}, {12}, {13}, {21}, {22}, {23}, {31}, {32}, {33}, as shown in figure 2. Now each person can distinguish between all states of the world.

Person 1 ●————————● Person 2

| 11 | 12 | 13 |
|----|----|----|
| 21 | 22 | 23 |
| 31 | 32 | 33 |

Person 1's partition

| 11 | 12 | 13 |
|----|----|----|
| 21 | 22 | 23 |
| 31 | 32 | 33 |

Person 2's partition

| r | r | r |
|---|---|---|
| r | r | s |
| s | s | s |

Person 1's actions

| r | r | s |
|---|---|---|
| r | r | s |
| r | s | s |

Person 2's actions
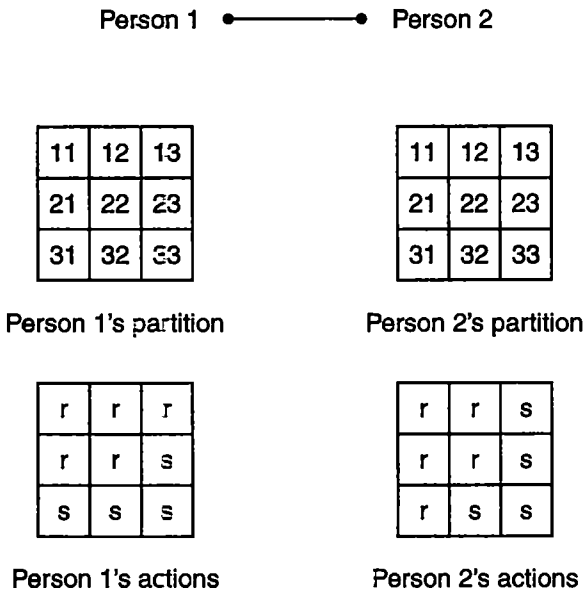
FIG. 2.—Complete network (full communication)

As before, person 1 revolts in states 11, 12, and 13 and stays at home in states 31, 32, and 33. Before, when person 1 had threshold 2, since he could not distinguish between states 21, 22, and 23, he chose to stay at home in these states. Now, person 1 can distinguish between these states and can take different actions in them. At state 21, person 1 revolts since he knows that person 2 has threshold 1 and will hence revolt. At state 23, person 1 stays at home since he knows that person 2 has threshold 3 and hence will not revolt. Similarly, person 2 revolts in state 12 and stays at home in state 32.

At state 22, when both people have threshold 2, person 1 wants to revolt if person 2 revolts, and person 2 wants to revolt if person 1 revolts. Hence, if both revolt, this is an equilibrium in the sense that each person is making the best choice given what the other person is doing (this is the standard concept of Nash equilibrium). It is also an equilibrium for both people to stay at home (if you do not revolt, I do not want to revolt, and if I do not revolt, you do not want to revolt). In our model, whenever there is this kind of indeterminacy (there is more than one equilibrium), we assume that the equilibrium that occurs is the one in which the most revolt takes place. That is, when two people with threshold 2 discover each other, we assume that they revolt.

To make a prediction for more general networks with $n$ people, it is
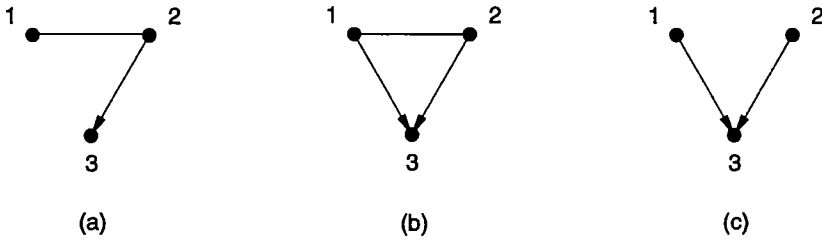
FIG. 3.—Three networks: everyone has threshold 2

simply a matter of writing down all states of the world, each person's information partition, and proceeding in a similar way. This becomes difficult, however, for large $n$ since there are $(n + 1)^n$ states of the world. But it is possible to solve the model without writing down all states of the world (of course, all but the simplest cases still require a computer; a program is available from the author). To take an example, consider the three networks in figure 3 and say everyone has threshold 2; thus, the true state of the world can be written 222.

First, consider network (a). Here, person 1 knows his own and person 2's threshold; thus, he knows that the state of the world is either 221, 222, 223, or 224. Person 2 knows her own and person 1's threshold and thus similarly knows that the state of the world is either 221, 222, 223, or 224. Person 3 knows his own and person 2's threshold and thus knows that the state of the world is either 122, 222, 322, or 422. Who revolts? If person 1 revolts at states 221, 222, 223, and 224, and person 2 revolts in states 221, 222, 223, and 224, this is an equilibrium since in all four states each person can count on the other to revolt. Will person 3 revolt? As far as person 3 can tell, the state of the world could be 422. In state 422, person 1 has threshold 4 and surely stays at home. In state 422, person 2 also stays at home because he cannot count on having a partner: person 1 stays at home, and person 2 does not know anything about person 3's threshold. So at state 422, person 3 will surely not have a partner. Since person 3 cannot distinguish between the states 122, 222, 322, or 422, he will not take the risk and hence will stay at home. So only persons 1 and 2 will revolt.

Note that here person 3 does not revolt even though he has threshold 2 and has a neighbor, person 2, who revolts. Person 3 knows that person 2 will revolt if person 1 is a willing companion, but person 3 does not know anything about person 1; as far as person 3 knows, person 1 could have threshold 4, in which case person 2 will surely stay at home.

Now, consider network (b). Again, persons 1 and 2 know that the state of the world is either 221, 222, 223, or 224. Now, person 3 knows every-
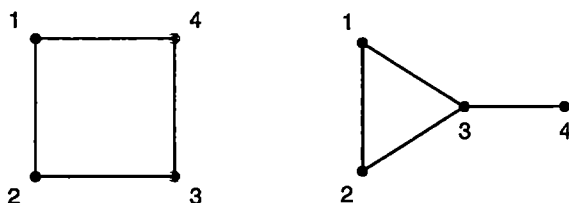
FIG. 4.—Square and kite

one's threshold and thus knows that the state of the world is exactly 222. Who revolts now? As before, person 1 and person 2 can count on each other and revolt in all four states 221, 222, 223, or 224. Person 3 knows that the state of the world is 222 and thus persons 1 and 2 will both revolt; hence, he revolts also. So now everyone revolts.

Finally, consider network (c). As before, person 3 knows everyone's threshold, but now person 1 only knows his own threshold and person 2 only knows her own threshold. As far as person 1 can tell, the state of the world could be any state in the set {211, 212, 213, 214, 221, . . . , 244}, and thus person 1 does not revolt. Similarly, person 2 does not revolt. Person 3 knows that the state is 222, but since persons 1 and 2 do not revolt at state 222, person 3 stays at home.

Note that person 3 revolts in network (b) but not in network (c), even though his knowledge of others' thresholds is the same in both cases. The reason is simply that in network (b) person 3 knows that persons 1 and 2 can communicate reciprocally and hence revolt, while in network (c), person 3 knows that they cannot. In other words, it is crucial here that person 3 knows whether persons 1 and 2 know each other; our model relies on the assumption that each person knows the relationships among her neighbors.

## PLURALISTIC IGNORANCE AND COMMON KNOWLEDGE

Now, say that we have four people, each with threshold 3, and consider two networks, the "square" and "kite," as shown in figure 4. These two networks nicely illustrate that a person's revolt depends not only on him knowing others' thresholds but also on him knowing what other people know.

Consider first the square. Since everyone has threshold 3, the actual state of the world can be written 3333. In the square, person 1 knows that person 2 and person 4 have threshold 3 but knows nothing about person 3's threshold. Hence, all person 1 knows is that the state of the world is in the set {3313, 3323, 3333, 3343, 3353}, where state 3353, for example,

is the state in which person 3 has threshold 5 and everyone else has threshold 3. Person 1 will revolt only if he knows that two other people will revolt in each of these states. Consider state 3353, which is perfectly possible as far as person 1 can tell. At state 3353, what, for example, will person 2 do? Person 2 does not know person 4's threshold and thus cannot distinguish between the states in the set {3351, 3352, 3353, 3354, 3355}. If person 2 revolts in these states, it is possible that she will have only one partner (the state could be 3355, in which case persons 3 and 4 never want to revolt). Hence, person 2 will stay at home at state 3353. So, if person 1 revolts in state 3353, at best only person 4 will join him. Hence, person 1 does not revolt at the states {3313, 3323, 3333, 3343, 3353} and thus does not revolt at state 3333. We can show similarly that persons 2, 3, and 4 do not revolt at state 3333 either.

Consider now the kite. Here, person 3 knows everyone's threshold and thus knows that the state of the world is 3333. Persons 1 and 2 know everything but person 4's threshold; they thus know the state of the world is in {3331, 3332, 3333, 3334, 3335}. Person 4 only knows his own and person 3's threshold; he thus knows that the state of the world is in the set {1133, 1233, . . . , 5533}. Person 4 clearly does not revolt; as far as he can tell, the state of the world is 5533, in which case he has at most one willing partner. But, if persons 1, 2, and 3 revolt in all states in {3331, 3332, 3333, 3334, 3335}, this is an equilibrium since in all three states at least three people revolt and each person has threshold 3.

So, according to our model, no one revolts in the square, but three people revolt in the kite. Note that this difference cannot be accounted for by macroscopic characteristics such as the total number of links (four symmetric links in both cases), or even by finer measures such as the number of neighbors each person has (in the kite, two revolters have only two neighbors, as in the square). The difference between the square and kite is truly a structural difference.

What is the difference between the square and the kite? In the square, each person knows that there are three people with threshold 3: himself and his two neighbors. That is, each person knows that conditions are such that revolt is possible. But, say I am considering whether to revolt. What do I know about, say, my neighbor to the right? I know that he has threshold 3. I also know that he knows that there is at least one other person with threshold 3: me. But I do not know anything about his other neighbor ("across" from me). Hence, I cannot count on him revolting, and so I do not revolt. In this case, even though everyone knows that revolt is possible, no one in fact revolts. In the kite, however, each person in the "triangle" knows not only that two other people have threshold 3, but they also know that these two other people know each other's threshold.

In other words, in the square, each person knows that there is sufficient

discontent (there are three people with threshold 3), but each person does not know that anyone else knows. It is not enough for everyone to simply know that there is sufficient discontent; what is required is "common knowledge" (Lewis 1969; Aumann 1976): everyone has to know that there is sufficient discontent, everyone has to know that everyone knows, everyone has to know that everyone knows that everyone knows, and so on. In the kite, among members of the triangle, the fact that there are three people with threshold 3 is common knowledge: each person knows it, knows that every other person knows it, and so on.

The importance of common knowledge for collective action can be demonstrated in a wide variety of contexts (starting with Lewis [1969]; see also, e.g., Rubinstein 1989; Chwe 1998). Several related ideas exist in the literature. One is the concept of "pluralistic ignorance" from social psychology (Allport 1924; Katz and Allport 1931; O'Gorman 1986), which refers to a situation in which people hold very incorrect beliefs about the beliefs of others. For example, in a 1972 survey, 15% of white Americans favored racial segregation, but 72% believed that a majority of the whites in their area favored segregation (O'Gorman 1979; see also Shamir 1993). Pluralistic ignorance is not usually understood as related to collective action; more often it is understood as a matter of individual psychology (O'Gorman 1986; Mullen and Hu 1988): a person reduces dissonance by thinking that her own view is the majority view, for example. Recently, however, it has been applied to the former Soviet Union and Eastern European states, the idea being that dissatisfaction was widespread but that few people knew how widespread it was. These accounts focus on limited communication due to criminal penalties for self-expression, a government-controlled press, and in the spirit of our model, a lack of social ties (Coser 1990; Kuran 1991). "The reduction of pluralistic ignorance," due to modern communication technology and increased foreign contacts, led to "a political wave of tremendous power" (Coser 1990, p. 182).

Another related idea is James Scott's (1990) distinction between the "public transcript," what subordinates say when talking publicly to their superiors, and the "hidden transcript," what subordinates say among themselves. For example, "if the sharecropping tenants of a large landowner are restive over higher rents, he would rather see them individually and perhaps make concessions than to have a public confrontation" (Scott 1990, p. 56). This is because public declarations enable subordinates to find out about each other. *It is only when this hidden transcript is openly declared that subordinates can fully recognize the full extent to which their claims, their dreams, their anger is shared by other subordinates* (Scott 1990, p. 223; emphasis in the original). The knowledge of others, which makes revolt possible, is formed not only by public declarations but also by social structure: "the informal networks of community . . . through

kinship, labor exchange, neighborhood, ritual practices, or daily occupational links" (Scott 1990, p. 151).

Even more explicitly, Roger Gould (1995, pp. 18–20) argues that "potential recruits to a social movement will only participate if they see themselves as part of a collectivity that is sufficiently large and solidary to assure some chance of success through mobilization. A significant source of the information they need to make this judgement is . . . social relations [which are] the mechanism for mutual recognition of shared interests (and of recognition of this recognition, and so on)." Translated into the language of our model, a person will participate only if she is assured that enough others will (the number of people revolting will be at least her threshold), and crucial in that decision is what she learns about the interests (thresholds) of others in her social network. Also, it is not just a matter of recognizing other people's interests (knowing other people's thresholds) but recognizing other people's recognition (knowing what others know), and so forth.

This idea, that social networks influence a group's knowledge of itself and thereby influence its ability to collectively act, is plausible and relevant enough to have arisen independently in multiple contexts. By giving it explicit mathematical form as in our model, it can be used not just to broadly motivate an argument but also to make specific predictions.

DYNAMICS

So far, our model is static. A simple way to add some dynamics is to assume that each person's neighborhood expands in time. We assumed earlier that a person's neighborhood is those people who are immediately adjacent, in other words, within one link away; we can think of this as the first "period" in time. We can make the simple assumption that in the second period each person's neighborhood is those people who are within two links away; in the third period, each person's neighborhood is those people within three links; and so on. To say this precisely, define the distance $d(j, i)$ from person $j$ to person $i$ to be the length of the shortest path from $j$ to $i$. Define $B(i, q) = \{j \in N: d(j, i) \leq q\}$ to be the "neighborhood of radius $q$" centered at $i$. The assumption is that at time $t$, person $i$'s neighborhood is $B(i, t)$: he knows the thresholds of everyone in $B(i, t)$ and the network relations among the people in $B(i, t)$. The idea here is that information flows through the network over time: first people know about their neighbors, second they learn about their neighbors' neighbors too, and so on. It is easy to show that as time progresses, and people learn more about each other, revolt never decreases: revolt either increases or remains constant.
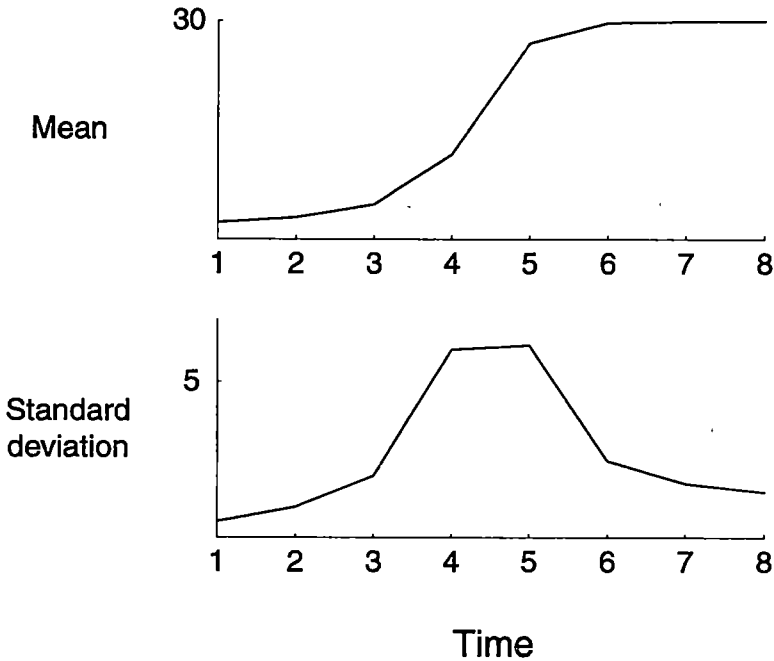
FIG. 5.—Mean and standard deviation of number of people revolting over time

## THRESHOLDS AND NETWORK POSITION

Since our model incorporates both network and threshold heterogeneity, it is natural to look at their combined effects. Consider networks of 30 people, in which each person has two neighbors selected at random; these are "unbiased" networks since any person is equally likely to be a given person's neighbor. Say that there are two people with threshold 1, two with threshold 2, and so on up to 15. We randomly generate 1,000 of these networks and find the average and standard deviation of the number of people revolting at a given time, as shown in figure 5.

The mean number of people revolting grows slowly at first, increases quickly, and finally slows down again as in the classic "logistic" curve. (Of course, in general in our model, depending on the thresholds and the network, the growth of revolt can be almost any curve, as shown in later examples; if growth is not logistic, one cannot conclude, as does, for example, Burt [1987], that the network has no effect.) By $t = 6$, in 958 of our 1,000 networks, all 30 people have revolted; by $t = 8$, in 994 of our 1,000 networks, all 30 have revolted. The standard deviation is highest in intermediate periods and can be quite substantial: for example, at $t = 4$, the

standard deviation is 6.0, more than half of the mean 11.6. This variance comes entirely from the randomness in network positions, since the distribution of thresholds is always the same; in other words, network variation alone can have a large effect on aggregate behavior.

One might expect that the people who revolt earliest are the ones with low thresholds. Figure 6 shows the threshold distribution of people who first revolt at time $t = 1$, $t = 2$, and so on; recall that there are two people for each threshold from 1 to 15. At $t = 1$, the two threshold 1 people both revolt, and occasionally a threshold 2 person revolts. At $t = 2$, the people who revolt are mainly threshold 2 and threshold 3 people; similarly, at time $t = 3$, the only people who revolt are those with low thresholds. At $t = 4$, some high-threshold people revolt, but the threshold distribution is still slightly skewed toward lower thresholds. At $t = 5$ and $t = 6$, the threshold distribution is skewed toward higher thresholds but is still fairly "flat." So it is true that people who revolt early tend to have low thresholds, but it is also true that people who revolt late do not necessarily have high thresholds. For example, when $t = 5$ and most of the people revolting have high thresholds, there are still many low-threshold people who are just starting to revolt.

Another way to see this is to plot the mean and standard deviation of the time at which a person of a given threshold starts to revolt, as in figure 7 (the rare cases in which a person does not revolt by $t = 8$ are not included in this calculation). People with higher thresholds indeed revolt later on average. However, except for people with threshold 1, the revolt time of people with low thresholds has higher variance than those with high thresholds. In other words, the revolt time of people with low thresholds, except of course people with threshold 1, are influenced much more by their network position. This makes sense: for people with high thresholds, all that matters for revolt is being connected to a large mass of people, but for people with low thresholds, whether you are lucky enough to be close to another person with a low threshold makes a big difference.

Empirical studies often make the assumption, implicit in linear regression, that all independent variables (age, social class, ideology, social position) enter linearly into participation. In other words, one's social position affects one's participation equally, regardless of whether one is very disposed to participate (has a low threshold) or is not disposed to participate (has a high threshold). This example suggests that this might not be justified: people with low thresholds would be affected by social position much more than people with high thresholds. In fact, perhaps another empirical strategy for verifying if social position is a significant influence on participation is to check whether the adoption time of people who seem to have low thresholds has greater variance than the adoption time of people who seem to have high thresholds.
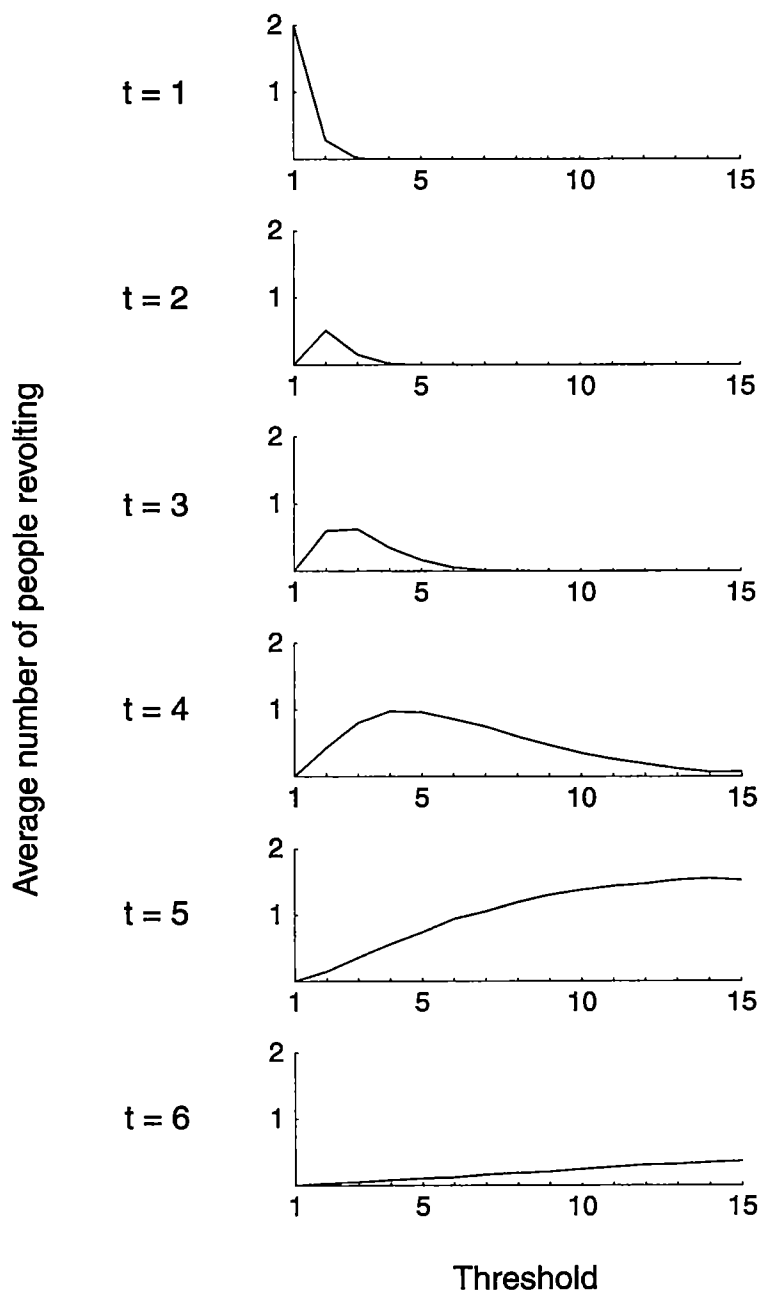
FIG. 6.—Threshold distribution of people who start to revolt at a given time
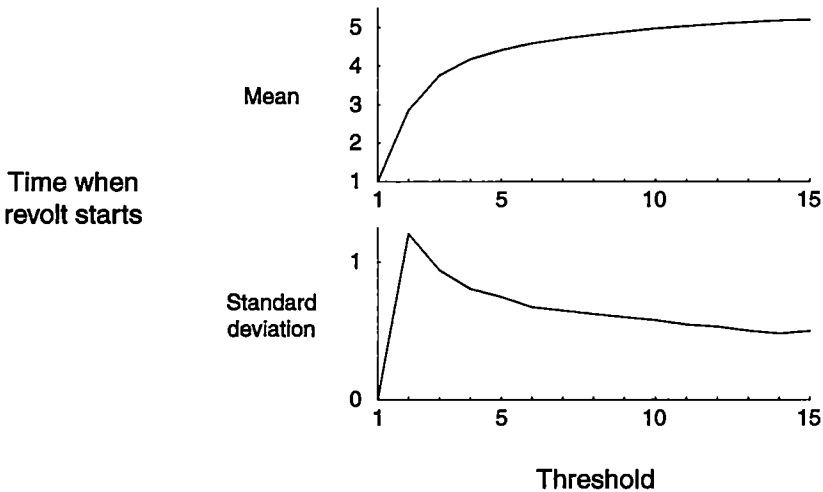
FIG. 7.—Mean and standard deviation of time period in which a person of a given threshold starts to revolt.

STRONG AND WEAK LINKS

The distinction between "strong" and "weak" links is an early insight of social network theory (Granovetter 1973). Roughly speaking, a strong link joins close friends and a weak link joins acquaintances. A simple empirical finding (Rapoport and Horvath 1961; see also White 1992) is that strong links tend to traverse a society "slowly": start with an arbitrary person, find two of her close friends, then find two close friends of each of these two people, and continue in this manner. As you iterate, the group increases slowly because often no one new is added: the close friends of my close friends tend to be my close friends also. If instead you successively add two acquaintances, the group grows quickly: the acquaintances of my acquaintances tend not to be my acquaintances. Weak links traverse a society "quickly": a demonstration suggests that any two people in the United States can be connected by as few as six weak links (Milgram 1992; see also Kochen 1989). Weak links tend to scatter widely, while strong links tend to be involuted.

To connect a large society, then, weak links are more important than strong links; weak links are more important for spreading information (Granovetter 1995; see also Montgomery 1991). For collective action, however, the importance of strong versus weak links is unclear. Data from volunteers in the 1964 Mississippi Freedom Summer, for example, show that the presence of a strong link to another potential participant correlates strongly and positively with participation, while the presence of a

141

weak link has no correlation (McAdam 1986; McAdam and Paulsen 1993; see also Fernandez and McAdam 1988). In three classic "diffusion" studies, rates of adoption are actually negatively correlated with the presence of weak links (Valente 1995, p. 51).

Our model shows how strong links can be better for participation when thresholds are low and weak links can be better when thresholds are high. This is because the involutedness of strong links, their tendency to form small cliques, is exactly what is needed to form common knowledge at a local level; if my friend's friend is my friend also, then common knowledge among the three of us is formed quickly. If thresholds are high, common knowledge must be formed among a large group of people; then, weak links are better simply because they speed up communication.

Again, consider networks of 30 people, each person having three neighbors. To randomly generate strong and weak networks, we use the following simple procedure (similar to that of Hedström [1994]; see, e.g., Fararo and Skvoretz [1987] and Frank and Strauss [1986] on biased networks in general and specifically Skvoretz [1990] on the difficulties of randomly generating biased networks). We first give each person a location, selected randomly and uniformly, on a unit square. A person's neighbors are selected in the following manner: with probability $p$, a "bias event" occurs and a neighbor is selected randomly among the closest three neighbors; with probability $1 - p$, a neighbor is selected randomly from the entire population. Hence, if $p = 1$, then each person's neighbors are simply the three closest people; if $p = 0$, then we have an unbiased random graph as discussed in the previous section.

Figure 8 shows some examples. When $p = 1$, the network is composed completely of strong links between geographically close people. When $p = 0.8$, on average 80% of the links are strong and 20% are weak links that scatter widely. When $p = 0.6$, there are more weak links, and when $p = 0$, locality is irrelevant and all links are weak.

For each of these four values of our "bias parameter" $p$, we generate 200 random networks. One way to check that these networks approximate what is normally considered strong versus weak is to plot a "tracing" for each value of $p$ (Rapoport and Horvath 1961), as shown in figure 9. Figure 9 shows the number of people within a certain number of links (averaged over all 30 people in a graph and over all 200 networks), for each of the four values of $p$. For example, for all $p$, starting from a given person, there are exactly four people within one link (the person herself and her three neighbors). The average number of people within two links of a given person (the person herself, her neighbors, and her neighbors' neighbors) is 11.3 for the $p = 0$ graphs and 6.7 for the $p = 1$ graphs. The lower $p$ is, that is, the more weak links in a network, the faster the tracing increases; weak-link networks are better at connecting large groups of people.
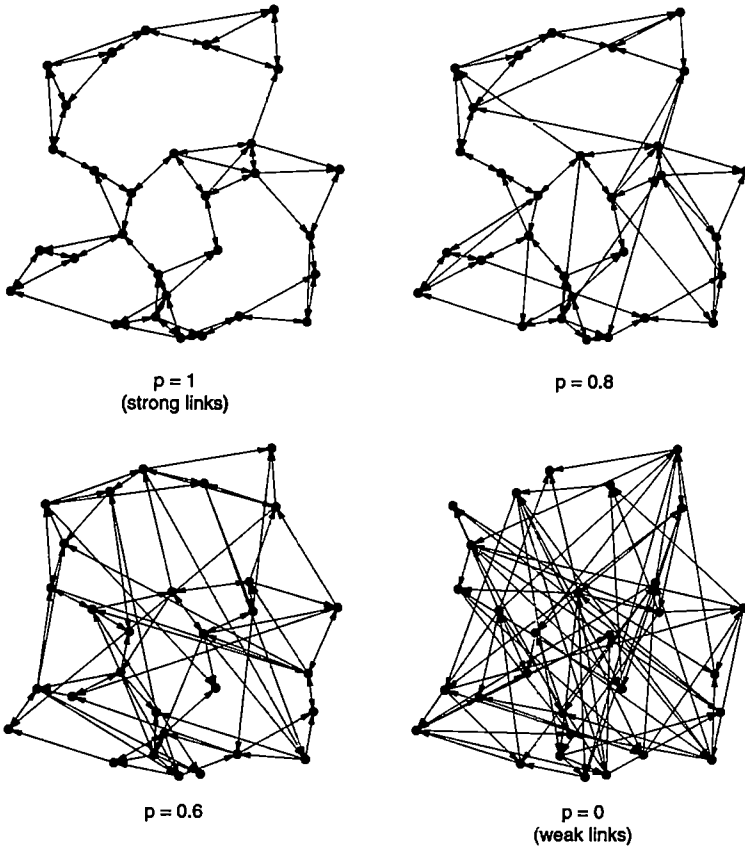
FIG 8.—Strong and weak links

Another way to see that our choices of $p$ correspond to strong versus weak is to calculate transitivity measures. Strong-link networks are supposed to be more transitive: in a strong-link network, it is more likely that a friend of a friend is my friend also. The simplest transitivity measure is just the number of transitive triads: the number of triples $(i, j, k)$ (where $i, j, k$ are all distinct) such that $i \rightarrow j$, $j \rightarrow k$, and $i \rightarrow k$. For the $p = 0$ weak-link networks, the number of transitive triads has mean 18.5 and standard deviation 4.3, and we can use this as a benchmark. The $p = 0.6$ networks have on average 36.0 transitive triads, roughly 4 standard deviations away from the benchmark, the $p = 0.8$ networks have on average 61.5 transitive triads, roughly 10 standard deviations away from the benchmark, and the $p = 1$ networks have on average 105.4 transitive triads, roughly 20 standard deviations away from the benchmark. Our
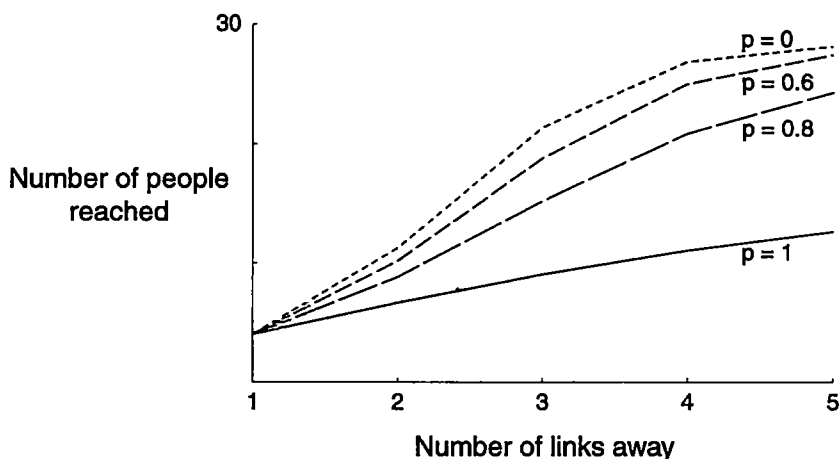
FIG. 9.—Tracings for $p = 1$, $p = 0.8$, $p = 0.6$, and $p = 0$

strong-link networks are real "outliers" when compared to the benchmark set of unbiased networks, but in this way are not completely unrealistic. For example, in their survey of 408 actual social networks, Holland and Leinhardt (1975) find the average number of transitive triples to be more than 5 standard deviations higher than what would be expected assuming a random "unbiased" distribution.

Assume that everyone has the same threshold $\theta$. When everyone has the same threshold, it is possible to show (Chwe 1999) that a person revolts at time $t$ if and only if his neighborhood contains a $t$-clique of size $\theta$ (a $t$-clique is a set of people in which each person is within $t$ links of every other). In other words, when everyone has the same threshold $\theta$, revolt is simply a matter of forming $t$-cliques of size $\theta$. Figure 10 shows the average number of people who revolt over time for each choice of $p$ and for $\theta$ ranging from 3 to 7. For example, in the first graph, everyone has threshold 3; here, the strong-link networks are very advantageous for revolt in the first time period. This is simply because the strong-link networks have many more cliques of size 3, as shown in figure 8. When everyone has threshold 4, the results are similar: strong links are advantageous (now in the second period); again, the transitivity or "involutedness" of strong-link networks helps form small cliques quickly.

When everyone has threshold 5, strong-link networks are better early on but worse in the long run. Here, strong-link networks form some cliques of size 5 quickly, but in the long run, the weak-link networks are better because of their greater "reach," as illustrated in the tracings in
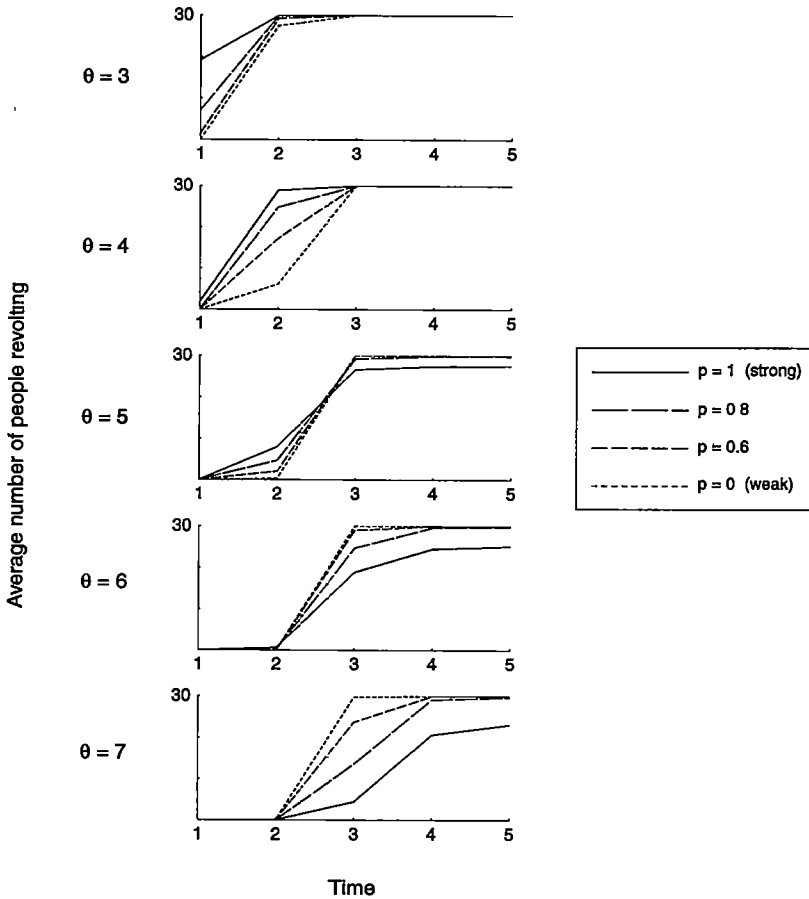
FIG. 10.—The number of people revolting when everyone has the same threshold θ, where θ = 3, 4, 5, 6, 7.

figure 9. When everyone has threshold 6 and everyone has threshold 7, weak-link networks start to clearly dominate.

That strong as opposed to weak links are important in recruiting for Freedom Summer is interpreted by McAdam (1986, p. 80) as the links having different functions: "although weak links may be more effective as diffusion channels, strong ties embody greater potential for influencing behavior." This is of course reasonable, but our model suggests it is not necessary. In Freedom Summer, a person might participate if a few friends also participate, which is like having a low threshold (strictly speaking, in our model only the number of fellow participants matters, not whether

they are close friends). If you and I are potential participants connected by a strong link, your friends are likely to be my friends, and participation among our group of friends would be common knowledge among us. If you and I are connected by a weak link, your friends and my friends do not know each other, and hence there is no common-knowledge group to which we both belong. In other words, the idea that weak links are always better for communication relies on the assumption that communication is about knowledge only and not "metaknowledge," knowledge of what others know. Strong links are better for forming common knowledge at a local level, and when thresholds are low, local mobilization is all that is necessary (see also Marwell, Oliver, and Prahl 1988, p. 532).

To summarize, the "structural" question of whether strong links or weak links are better for collective action cannot be answered without considering the "strategic" situation, the distribution of thresholds. Structural and strategic considerations interact in interesting ways.

## LIMITED COMMUNICATION AND THE FRAGILITY OF COLLECTIVE ACTION

One of the earliest and simplest explicit models of a coordination process was introduced by Schelling (1978): first, only people with low thresholds participate, but their participation makes people with slightly higher thresholds want to participate. As the number participating grows, people join in successively, in a "snowball" or "bandwagon" effect. I call this the "simple model."

This simple model can be represented as a special case of our model; our model might thus be considered its generalization to arbitrary networks. Doing this shows that the simple model implicitly assumes that communication possibilities are extremely restricted: a person only gets information about people who have lower thresholds, making bilateral communication impossible. One compelling result of the simple model is that collective action depends crucially on the thresholds of people "early" in its growth (Granovetter 1978). I show that this fragility depends heavily on the assumption of extremely restricted communication.

To take an example, say that there are five people: one with threshold 1, one with threshold 2, and so on up to threshold 5. The simple model goes like this: first, the threshold 1 person willingly revolts by himself. When the threshold 2 person sees one person revolting, that is enough to make him want to revolt, and he joins in. Then the threshold 3 person joins in, and so forth until everyone joins in. Note that in this process, there is no explicit communication and indeed no need to learn anyone else's threshold: each person makes their own independent decision given the number of people already revolting.

146
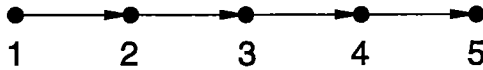
FIG. 11.—"Bandwagon"

To put this in our framework, one simply considers the network shown in figure 11, where the numbers 1, 2, 3, 4, 5 indicate each person's threshold. Our model has quite different premises: people do not observe each other's actions but learn their thresholds. Still, the two models give the same conclusion. In the first period, the threshold 1 person naturally revolts and the threshold 2 person revolts because he knows that there is a threshold 1 person. In the second period, the threshold 3 person revolts because he learns that there is a threshold 1 person and a threshold 2 person, and so on.

Note that the communication network that adapts our model to the simple model is very sparse; in fact, it could not be made any more sparse without disconnecting the group. Communication is ample in that a person eventually learns about everyone with a lower threshold (Braun 1995), but restricted in that a person never learns about anyone with a higher threshold; communication is in only one direction, never reciprocal.

The effects of this sparse network are clear in an example. Start with the same network and the same five people. Now change the first person's threshold from 1 to 2 so thresholds are now 2, 2, 3, 4, 5. Now, no one ever revolts. The first person no longer wants to jump in all by himself, and hence the second person does not follow. Participation is vulnerable to small perturbations of people "early" in the bandwagon.

However, if the first and the second person could talk reciprocally, as in the network in figure 12, the two would jump in together and the bandwagon would go on as before. In other words, the breakdown of revolt when going from 1, 2, 3, 4, 5 to 2, 2, 3, 4, 5 depends entirely on the assumption that communication is extremely limited, never reciprocal.

To illustrate this more generally, again consider networks of 30 people in which each person has three neighbors. Again, say two people have threshold 1, two people have threshold 2, and so on up to 15. To construct the social network, we first place the people in a line, ordered by their threshold, like this: 1, 1, 2, 2, 3, 3, . . . , 15, 15. A person's neighbors are selected in the following manner: with probability $q$, a "bias" or "skew"
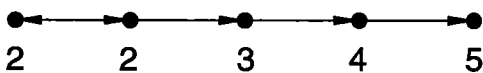


FIG. 12.—If the network is less sparse, the bandwagon is less fragile

147

event occurs and a neighbor is selected randomly from a person to the left (if there is no one to the left, no one is selected); with probability $1 - q$, a neighbor is selected randomly from the entire population. Hence, if $q = 1$, communication only flows in one direction, from left to right, as in figure 11. If $q = 0$, we have an unbiased random graph, as discussed earlier.

Say we set $q = 1$ and generate a "maximally skewed" random graph in this manner. Revolt then starts with the two threshold 1 people and then spreads to people with higher thresholds. However, if we perturb the situation by letting the first two people have threshold 2 instead of threshold 1 (in other words, instead of thresholds being 1, 1, 2, 2, 3, 3, ..., 15, 15, they are 2, 2, 2, 2, 3, 3, ..., 15, 15), revolt collapses completely, as in the simple model. Again, this is because when communication is never reciprocal, threshold 1 people are necessary to make any revolt at all possible.

What happens as we move $q$ away from 1, so our networks are not completely skewed? Let $q = 1, 0.9, 0.8, 0.7, 0.6, 0$, and randomly generate 200 sample graphs for each value of $q$. Figure 13 shows the average number of people revolting over time, where the black line corresponds to the original thresholds 1, 1, 2, 2, 3, 3, ..., 15, 15 and the gray line corresponds to the perturbed thresholds 2, 2, 2, 2, 3, 3, ..., 15, 15. Note that for all values of $q$, revolt in the first period is much greater with the original than with the perturbed thresholds; this is simply because the original thresholds include two people with threshold 1. When $q = 1$, with the perturbed thresholds, revolt completely collapses. When $q = 0.9$, however, by the third period, revolt with the perturbed thresholds is half or more of that of the original thresholds. As $q$ decreases from $q = 0.8$, the perturbation quickly makes relatively little difference. Another interesting thing to note is that there is an "optimal" level of $q$; for both the original and perturbed thresholds, revolt increases faster when $q = 0.6$, for example, than when either $q = 0$ or $q = 1$. In other words, there is an advantage to having information flow from people with lower thresholds to people with higher thresholds, but too much skew is bad because it reduces reciprocal communication.

To summarize, the simple model greatly overestimates the fragility of collective action; even when only 10% or 20% of the links are random and not "skewed," revolt does not collapse with a small perturbation of thresholds. Putting it another way, the less skewed the network and the more possibilities for reciprocal communication, the less fragile is revolt.

One can defend the simple model as applying to a spontaneous, unorganized process, explicit communication being unnecessary or unavailable. Communication is not necessarily automatic or easy, but it is hard to believe that a few words ("I'll do it if you do it") or even eye contact between
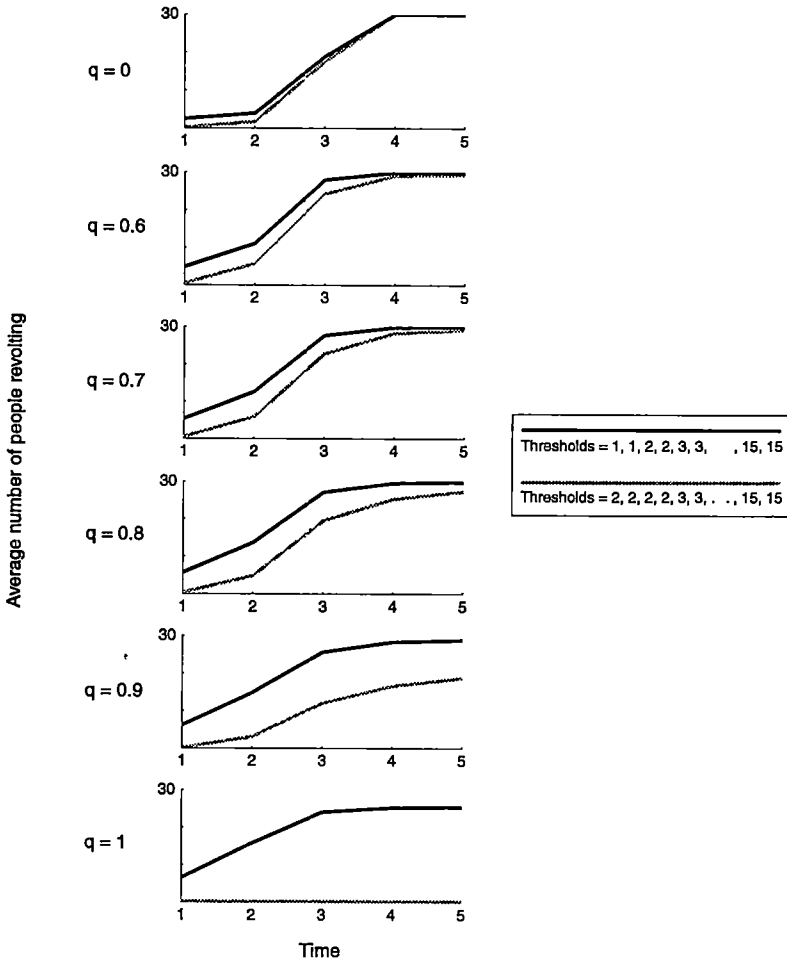
FIG. 13.—If the network is less skewed, the bandwagon is less fragile

two people at least could not arise spontaneously. "Research on participation in a variety of prosaic gatherings, in political, religious, and sports demonstrations, and probably in riots, establishes that most participants are neither alone or anonymous; rather they assemble with or soon encounter family members, friends, or acquaintances" (McPhail 1991, p. 14). In any case, the simple model does not approximate the "general" case; when coordination takes place under the simple model, it takes place under communicative conditions so limited that no pair of individuals can reciprocally communicate.

## CONCLUDING REMARKS

The broad message of this article is that there is no necessary opposition between structure and rationality; our model incorporates both explicit network structure and complete rationality and derives results not obtainable from either perspective alone. This supposed opposition, which has a long and undoubtedly complex historical basis, is evident in current research: almost all models of the effects of social structure on individual action make adaptive, bounded rationality, or behavioral assumptions. This is true not only in sociology (e.g., Gould 1993; Macy 1991, 1995; but see also Marwell and Oliver 1993) but also in recent work in economics (e.g., Akerlof 1997; Anderlini and Ianni 1996; Blume 1993, 1995; Ellison 1993; Goyal 1996; Mailath, Samuelson, and Shaked 1997; Morris 1999; Temzelides 1997; Tesfatsion 1997; Young 1996, 1998). Sometimes this supposed opposition is drawn very sharply; for example, in his discussion of the Schelling-Granovetter threshold model, Michael Macy (1991, p. 735) finds that "structural analysis is badly handicapped by Granovetter's rational choice assumption . . . that individual participation is triggered by the contribution rate of the group as a whole. The choices of those in one's immediate circle have no special relevance. . . . The structure of social ties is much more relevant if the actions of others shape individual behavior directly rather than as a by-product of an underlying interest in marginal utility." Here, adding structure seems to require the immediate jettisoning of rationality, as opposed to any other assumption (see, e.g., Krassa 1988; Valente 1996), and seems to lead directly and unarguably toward a model in which people do not make choices but rather are influenced directly by others.

One advantage of explicitly modeling individual choices, as opposed to letting behavior directly affect behavior, is that one is led to recognize the wide variety of strategic situations in which networks play a role and to develop models that are sensitive to this variety. For example, Hedström (1994, p. 1162) develops a model in which a person's probability of participating is a linear function, weighted by social distance, of the participation of others and readily admits that, for the sake of simplicity, the model "assumes that the decision to join an organization for collective action resembles the decision whether or not to buy a private good." But surely these are quite different decisions; the entire theory of public goods and the "free-rider problem," for example, depends on this distinction. Instead of effacing this distinction, a network model could more interestingly interact with it: networks that are good for the diffusion of private goods purchases might be quite bad for the spread of collective political action, for example.

Similarly, what exactly people do with the network can also be made

explicit. It is commonly argued that a social network is used to communicate information, but information is almost never modeled explicitly (the "state space" model, used here, is one of several ways to model it). Also, it is often unclear what the information is about. As in the Schelling-Granovetter model, a person usually is assumed to react to the actions of her neighbors in the network. In my model, a person learns about her neighbors' preferences, their willingness to revolt, and does not directly respond to their actions. This distinction is not merely conceptual but can be found in "real life," for example, in Dingxin Zhao's (1998) analysis of the April 27, 1989, student demonstration in Beijing. A crucial "ecological" factor for participation was that students lived close together in university dormitories, allowing them to discuss and "share their anger." This communication process, which enabled students to know that enough others felt the same way, is what my model tries to describe. Once students marched on the street, however, students on bicycles rode back and forth between separate groups, thereby informing each group that other groups were already marching, and there was a "snowball" effect as in the Schelling-Granovetter model.

One specific contribution of my model is that it allows a person's choice to depend on the relationships among his neighbors. I make the strong assumption that a person knows whether two of his neighbors know each other, and as illustrated in the many examples above, what a person knows about the knowledge of other people is crucial. The extent to which this assumption is empirically valid is not yet settled: some empirical researchers (notably Bernard, Kilworth, and Seiler 1980, 1982) find that a person's knowledge of specific instances of social interactions among the people around her is typically quite poor, while others (e.g., Freeman, Romney, and Freeman 1987; Freeman, Freeman, and Michaelson 1988; see also Romney and Faust 1982) find that a person is typically fairly good at understanding the long-term patterns of interaction around her. Presumably, one reason this question is interesting is because it might affect individual and thus social behavior. However, in almost all existing models, including models in which a person's action is a linear function of the actions of his neighbors (e.g., Gould 1993; Hedström 1994), learning models (e.g., Macy 1991), and models in which people can explicitly organize each other and sign all-or-nothing contracts (Marwell and Oliver 1993), the question is irrelevant. In all of these models, a person's action depends only on her neighbors and their actions and not on the connections among her neighbors; connections among neighbors do affect the long-run evolution of society-wide behavior but do not affect an individual's action at a given moment. Whether the two friends I get information from know each other or not should have an effect on my decision. Although this effect depends on me knowing whether my two friends know

each other, it is also unwarranted to a priori rule out this effect completely, as do existing models.

In this way, existing models are somewhat limited in looking at network transitivity; for example, cliques are often found to be bad for collective action, despite the intuition that collective action starts in small "subcultures," which are commonly thought of and "detected" as cliques. To take another example, there is some evidence that people's beliefs are systematically biased toward transitivity: people tend to think that the networks they belong to are more transitive than they really are (see, e.g., Freeman 1992). To explore the implications of this bias for individual and social action, one needs a model in which transitivity directly matters. Finally, there is even an immediate implication for survey methods: when Opp and Gern (1993), for example, surveyed participants in the demonstrations that led to the collapse of East Germany, they asked each person whether he had friends who participated and found that this was a significant variable in predicting his participation. Our model suggests that each person should also be asked if his friends who participated knew each other. If this variable is significant (of course after controlling for spurious correlations), then this would be evidence for a communication model as in this article.

In sum, structure and strategy can work together in various interesting and unexpected ways: for example, the game-theoretic model here, in which knowledge of other people's knowledge is crucial, allows transitivity to affect individual action more directly than bounded rationality or behavioral models seemingly more congenial to structural considerations. The structural approach and the strategic approach to collective action can surely find more insights together than separately.

## APPENDIX

Here I specify explicitly people's utilities, knowledge, and strategies, and then consider equilibria. First, define utility: given person $i$'s own threshold $\theta_i$ and everyone's actions $a = (a_1, \ldots, a_n)$, her utility $u_i(\theta_i, a_1, \ldots, a_n)$ is given by

$$u_i(\theta_i, a_1, \ldots, a_n) = \begin{cases} 0 & \text{if } a_i = s \\ 1 & \text{if } a_i = r \text{ and } \#\{j \in N : a_j = r\} \geq \theta_i \\ -z & \text{if } a_i = r \text{ and } \#\{j \in N : a_j = r\} < \theta_i, \end{cases}$$

where $-z < -(n + 1)^n$. In other words, a person always gets utility 0 by staying at home. If he revolts, he gets utility 1 if the total number of people revolting is at least $\theta_i$, his threshold. If he revolts and not enough people join him, he gets the very large penalty $-z$.

152

I represent person $i$'s knowledge as a partition of the set of all possible states of the world. Since the uncertainty here is over everyone's thresholds, the set of all possible states of the world is $\Theta = \{1, 2, 3, \ldots, n + 1\}^n$. For simplicity, I assume that thresholds are independently and identically distributed, and that each threshold is equally likely: thus the probability distribution $\pi:\Theta \to [0,1]$ over $\Theta$ is given by $\pi(\theta) = 1/(n + 1)^n$. Since people are also assumed to have incomplete information about the network (person $i$ only knows the network structure within his neighborhood $B(i)$), I should also explicitly model uncertainty in the network. However, given our other particular assumptions, doing this would not change any of the results.

Person $i$ knows the thresholds of people in the neighborhood $B(i)$. Hence, if the actual state of the world is $\theta \in \Theta$, then person $i$ knows only that the state of the world is in the set $P_i(\theta) = \{(\theta_{B(i)}, \phi_{N-B(i)}): \phi_{N-B(i)} \in \{1, 2, 3, \ldots, n + 1\}^{n - \#B(i)}\}$ (we use the notation $\theta_A = (\theta_j)_{j \in A}$). Taken together, the sets $\{P_i(\theta)\}_{\theta \in \Theta}$ form a partition of $\Theta$, which we call $\mathcal{P}_i$.

Person $i$'s strategy is a choice of action given her knowledge of the state of the world. Thus, I define a strategy for person $i$ to be a function $f_i$: $\Theta \to \{r, s\}$ which is measurable with respect to $\mathcal{P}_i$, that is, for all $\theta, \theta' \in \Theta$, if $\theta, \theta' \in P$, where $P \in \mathcal{P}_i$, then $f_i(\theta) = f_i(\theta')$. The idea here is that if $\theta$ and $\theta'$ are in the same element of the partition $\mathcal{P}_i$, then person $i$ cannot distinguish between the two states $\theta$ and $\theta'$, and hence must take the same action in the two states. Say that $F_i$ is the set of all strategies of person $i$.

Given strategies $(f_1, \ldots, f_n)$, person $i$'s expected utility is $EU_i(f_1, \ldots, f_n) = \sum_{\theta \in \Theta} \pi(\theta)u_i(\theta_i, f_1(\theta), \ldots, f_n(\theta))$. We say that $(f_1, \ldots, f_n)$ is an equilibrium if for all $i \in N$, and for all $g_i \in F_i$, $EU_i(f_1, \ldots, f_n) \geq EU_i(f_1, \ldots, f_{i-1}, g_i, f_{i+1}, \ldots, f_n)$. In other words, $(f_1, \ldots, f_n)$ is an equilibrium if no individual can gain by deviating to another strategy $g_i \in F_i$.

In general, there exist many equilibria. It is not hard to show that there uniquely exists an equilibrium $(f_1^*, \ldots, f_n^*)$ such that if $(f_1, \ldots, f_n)$ is an equilibrium, then $f_i(\theta) = r \Rightarrow f_i^*(\theta) = r$. In other words, the equilibrium $(f_1^*, \ldots, f_n^*)$ has the greatest possible revolt: if person $i$ revolts in any equilibrium, she revolts in this "best" equilibrium. We assume that this equilibrium is the one which is played. Given the relation $\to$ and thresholds $\theta$, the set of people who revolt is $\{i \in N: f_i^*(\theta) = r\}$.

REFERENCES

Akerlof, George A. 1997 "Social Distance and Social Decisions." *Econometrica* 65: 1005–27.
Allport, Floyd H. 1924. *Social Psychology*. Boston: Houghton Mifflin.
Anderlini, Luca, and Antonella Ianni. 1996. "Path Dependence and Learning from Neighbors." *Games and Economic Behavior* 13:141–77.

Aumann, Robert J. 1976. "Agreeing to Disagree." *Annals of Statistics* 4:1236–39.

Bernard, H. Russell, Peter D. Kilworth, and Lee Sailer. 1980. "Informant Accuracy in Social-Network Data IV: A Comparison of Clique-Level Structure in Behavioral and Cognitive Network Data." *Social Networks* 2:191–218.

———. 1982. "Informant Accuracy in Social-Network Data V. An Experimental Attempt to Predict Actual Communication from Recall Data." *Social Science Research* 11:30–66.

Blume, Larry. 1993. "The Statistical Mechanics of Strategic Interaction." *Games and Economic Behavior* 5:387–424.

———. 1995. "The Statistical Mechanics of Best-Response Strategy Revision." *Games and Economic Behavior* 11:111–45.

Braun, Norman. 1995. "Individual Thresholds and Social Diffusion." *Rationality and Society* 7:167–82.

Burt, Ronald S. 1987 "Social Contagion and Innovation: Cohesion versus Structural Equivalence." *American Journal of Sociology* 92:1287–335.

Chwe, Michael Suk-Young. 1998. "Culture, Circles, and Commercials: Publicity, Common Knowledge, and Social Coordination." *Rationality and Society* 10:47–75.

———. 1999 "Communication and Coordination in Social Networks." *Review of Economic Studies*, in press.

Coser, Lewis A. 1990. "The Intellectuals in Soviet Reform: On 'Pluralistic Ignorance' and Mass Communications." *Dissent* spring:181–83.

Ellison, Glenn. 1993. "Learning, Local Interaction, and Coordination." *Econometrica* 61:1047–71.

Fararo, Thomas J., and John Skvoretz. 1987 "Unification Research Programs: Integrating Two Structural Theories." *American Journal of Sociology* 92:1183–1209.

Fernandez, Roberto M., and Doug McAdam. 1988. "Social Networks and Social Movements: Multiorganizational Fields and Recruitment to Freedom Summer." *Sociological Forum* 3:357–82.

Frank, Ove, and David Strauss. 1986. "Markov Graphs." *Journal of the American Statistical Association* 81:832–42.

Freeman, Linton C. 1992. "Filling in the Blanks: A Theory of Cognitive Categories and the Structure of Social Affiliation." *Social Psychology Quarterly* 55:118–27.

Freeman, Linton C., Sue C. Freeman, and Alaina G. Michaelson. 1988. "On Human Social Intelligence." *Journal of Social and Biological Structures* 11:415–25.

Freeman, Linton C., A Kimball Romney, and Sue C. Freeman. 1987. "Cognitive Structure and Informant Accuracy." *American Anthropologist* 89:310–25.

Gould, Roger V. 1993. "Collective Action and Network Structure." *American Sociological Review* 58:182–96.

———. 1995. *Insurgent Identities: Class, Community, and Protest in Paris from 1848 to the Commune.* Chicago: University of Chicago Press.

Goyal, Sanjeev. 1996. "Interaction Structure and Social Change." *Journal of Institutional and Theoretical Economics* 152:472–94.

Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360–80.

———. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83:1420–43.

———. 1995. *Getting a Job: A Study of Contacts and Careers*, 2d ed Chicago: University of Chicago Press.

Hedström, Peter. 1994. "Contagious Collectivities· On the Spatial Diffusion of Swedish Trade Unions, 1890–1940." *American Journal of Sociology* 99:1157–79.

Holland, P. W., and S. Leinhardt. 1975. "The Statistical Analysis of Local Structure in Social Networks." Pp. 1–45 in *Sociological Methodology 1976*, edited by D. R. Heise. San Francisco· Jossey-Bass.

Katz, Daniel, and Floyd H. Allport. 1931. *Student Attitudes*. Syracuse, N.Y.: Craftsman Press.

Kochen, Manfred. 1989. *The Small World*. Norwood, N J.: Ablex Publishing.

Krassa, Michael A 1988. "Social Groups, Selective Perception, and Behavioral Contagion in Public Opinion." *Social Networks* 10:109–36.

Kuran, Timur. 1991. "Now Out of Never: The Element of Surprise in the East European Revolution of 1989." *World Politics* 44:7–48.

Lewis, David K. 1969. *Convention: A Philosophical Study*. Cambridge, Mass.: Harvard University Press.

Macy, Michael W. 1991. "Chains of Cooperation: Threshold Effects in Collective Action." *American Sociological Review* 56·730–47.

————. 1995. "Beyond Rationality in Models of Choice." *Annual Review of Sociology* 21:73–91.

Mailath, George J., Larry Samuelson, and Avner Shaked. 1997. "Correlated Equilibria and Local Interactions." *Economic Theory* 9:551–56.

Marwell, Gerald, and Pamela Oliver. 1993. *The Critical Mass in Collective Action*. Cambridge: Cambridge University Press

Marwell, Gerald, Pamela E. Oliver, and Ralph Prahl. 1988. "Social Networks and Collective Action: A Theory of the Critical Mass. III." *American Journal of Sociology* 94:502–34.

McAdam, Doug. 1986. "Recruitment to High-Risk Activism: The Case of Freedom Summer." *American Journal of Sociology* 92:64–90.

McAdam, Doug, and Ronnelle Paulsen. 1993. "Specifying the Relationship between Social Ties and Activism " *American Journal of Sociology* 99:640–67.

McPhail, Clark. 1991. *The Myth of the Madding Crowd*. New York: Aldine de Gruyter.

Milgram, Stanley. 1992. *The Individual in a Social World: Essays and Experiments*, 2d ed. Edited by John Sabini and Maury Silver New York: McGraw Hill.

Montgomery, James D. 1991. "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis." *American Economic Review* 81:1408–18.

Morris, Stephen. 1999. "Contagion." *Review of Economic Studies*, in press.

Mullen, Brian, and Li-tze Hu. 1988. "Social Projection as a Function of Cognitive Mechanisms: Two Meta-Analytic Integrations." *British Journal of Social Psychology* 27·333–56.

O'Gorman, Hubert J. 1979. "White and Black Perceptions of Racial Values." *Public Opinion Quarterly* 43:48–59.

————. 1986. "The Discovery of Pluralistic Ignorance: An Ironic Lesson." *Journal of the History of the Behavioral Sciences* 22:333–47.

Opp, Karl-Dieter, and Christiane Gern. 1993. "Dissident Groups, Personal Networks, and Spontaneous Cooperation " *American Sociological Review* 58.659–80.

Rapoport, Anatol, and W. J. Horvath. 1961. "A Study of a Large Sociogram." *Behavioral Science* 6.279–91.

Romney, A Kimball, and Katherine Faust. 1982. "Predicting the Structure of a Communications Network from Recalled Data." *Social Networks* 4:285–304.

Rubinstein, Ariel. 1989. "The Electronic Mail Game: Strategic Behavior under 'Almost Common Knowledge.' " *American Economic Review* 79:385–91.

Schelling, Thomas C. 1978. *Micromotives and Macrobehavior*. New York: Norton.

Scott, James C. 1990 *Domination and the Arts of Resistance: Hidden Transcripts*. New Haven, Conn.: Yale University Press.

Shamir, Jacob. 1993. "Pluralistic Ignorance Revisited: Perception of Opinion Distributions in Israel." *International Journal of Public Opinion Research* 5:22–39.

Skvoretz, John. 1990. "Biased Net Theory: Approximations, Simulations, and Observations." *Social Networks* 12:217–38.

Temzelides, Ted. 1997. "Evolution, Coordination, and Banking Panics." *Journal of Monetary Economics*.

Tesfatsion, Leigh. 1997. "How Economists Can Get a Life." Pp. 533–64 in *The Economy as an Evolving Complex System II*, edited by W. Brian Arthur, Steven N. Durlauf, and David A. Lane. Redwood City, Calif.: Addison-Wesley Longman.

Valente, Thomas W. 1995. *Network Models of the Diffusion of Innovations*. Creskill, N.J.: Hampton Press.

———. 1996 "Social Network Thresholds in the Diffusion of Innovations." *Social Networks* 18:69–89.

White, Harrison C. 1992. *Identity and Control: A Structural Theory of Social Action*. Princeton, N.J.: Princeton University Press

Young, Peyton. 1996. "The Economics of Convention." *Journal of Economic Literature* 10:105–22.

———. 1998. *Individual Strategy and Social Structure : An Evolutionary Theory of Institutions*. Princeton, N.J.: Princeton University Press.

Zhao, Dingxin. 1998. "Ecologies of Social Movements: Student Mobilization during the 1989 Prodemocracy Movement in Beijing." *American Journal of Sociology* 103: 1493–529.

# Interracial Conflict and Interracial Homicide: Do Political and Economic Rivalries Explain White Killings of Blacks or Black Killings of Whites?[1]

David Jacobs
*Ohio State University*

Katherine Wood
*University of Chicago*

What factors lead to interracial killings? Because racial conflict explanations have been overlooked in the previous literature, this article studies the determinants of disaggregated interracial killing rates in 165 U.S. cities by testing economic, political, and social control accounts. After holding the probability of interracial contacts and the total murder rate constant, the results show that cities with a black mayor and greater economic competition between the races have more white killings of blacks. The same hypotheses explain black killings of whites, but these killings are less likely in cities with black mayors. Police department size does not explain white killings of blacks, but cities with larger departments have fewer black killings of whites. The findings suggest that economic rivalries and contests for political influence lead to greater interracial violence.

Interracial killings represent some of the most intense and violent conflicts between members of groups that have not been treated equally. When the determinants of these lethal events were investigated, most researchers

ignored racial antagonisms. It is difficult to believe that the animosities created by economic and political struggles between blacks and whites do not explain at least some of the variation in interracial homicides. Despite the interest in explaining crime with theories about group conflict (Turk 1969; Taylor, Walton, and Young 1973), few researchers have convincingly applied this perspective to interracial violent crime. Of course, not all interracial killings can be attributed to racial antagonisms. This study nevertheless focuses on economic and political rivalries between the races to see if such conflicts lead to more black killings of whites or white killings of blacks.

A study of such relationships should tell us more about the political content of interracial violence. Many neo-Marxists and some Weberian conflict theorists see crime as inarticulate political resistance against unequal social arrangements. Yet victimization patterns are an embarrassment for attempts to explain crime with hypotheses about racial or class conflict because the victims of crime are most likely to be poor and minorities. Conflict theorists who try to account for crime with hypotheses that treat law breaking as embryonic rebellion must explain why most crimes are not directed against the dominant race or class. A study that focuses on the political and economic divisions that may lead to interracial crimes is uniquely suited to assess the explanatory power of such political accounts. By focusing only on the killings that involve offenders and victims from the most dominant and the most subordinate race, we can isolate relationships between racial conflicts over scarce resources and lethal violence between blacks and whites.

Some explanations for crime emphasize the social arrangements that produce crime. We derive hypotheses about the social conditions that lead to interracial killings from the minority-group threat theories that were developed to explain prejudice, discrimination, and interracial conflict. These hypotheses are assessed by looking at economic and political threats to majority group dominance that may generate interracial violence. Many studies show that threats produced by interracial economic competition lead to enhanced interracial aggression (Beck and Tolnay 1990; Olzak 1992; Olzak, Shanahan, and McEneaney 1996) and prejudice (Quillian 1995, 1996), yet such explanations have not been used successfully to explain interracial homicides. We extend this study of competitive processes by examining political rivalries between blacks and whites to see if these overlooked rivalries explain interracial killings.

Alternative explanations for crime stress social control. The possibility that the police can suppress some forms of interracial crime has not been carefully investigated. The procedures the police must use may inhibit their ability to control some interracial killings but not others. We investi-

gate such control hypotheses by seeing if larger police departments produce fewer black killings of whites or fewer white killings of blacks.

We focus on interracial killings for two reasons: First, data on these events are far more accurate than data about other crimes. Second, the alternative interracial crimes differ so much from homicides that they should be the subject of separate analyses. Because interracial homicides involving other races are infrequent, we limit this investigation to white killings of blacks and black killings of whites. We analyze these two interracial homicide rates separately because models that best explain each rate should not be the same. Most studies of interracial homicides use total interracial killing rates (see Messner and Golden [1992] and Messner and South [1992] for exceptions). Yet an analysis of the combined rates rests on the dubious assumption that coefficients on the determinants of black killings of whites are identical to the coefficients on the determinants of white killings of blacks.

Researchers probably combined these dissimilar dependent variables for methodological reasons. Interracial killings are infrequent events. Some cities have no black killings of whites or white killings of blacks. Ordinary least squares (OLS) models give inconsistent estimates when the dependent variable is censored in this way, but there is a statistical procedure that handles censoring. We estimate the models with weighted Tobit. We also average the interracial killing rates over a seven-year period to reduce instability and the number of cities with no interracial killings. Finally, because inclusive research designs are most likely to produce unbiased estimates (Johnston 1984; see also n. 10 below), we assess many hypotheses, but this means that the theoretical section cannot focus on just a few explanations.

THE LITERATURE AND SOME HYPOTHESES

We present explanations for interracial killings from three perspectives. We first present hypotheses based on heterogeneity theory (Blau 1977), which has provided the theoretical basis for almost all of the research on interracial killings. We then use the minority-group threat literature (Blumer 1958; Blalock 1967; Quillian 1995, 1996) that stresses interracial economic (Beck and Tolnay 1990; Olzak 1992) and political competition (Wilson 1978; Bobo and Gilliam 1990) to derive conflict hypotheses about white killings of blacks or black killings of whites. The last part of the theoretical discussion develops contrasting hypotheses about social control. Most white killings of blacks result from interpersonal disputes, but black killings of whites are more likely to occur during a felony. This contrast suggests that police size will lead to reduced black killings of

whites, yet the relative size of departments may not influence white killings of blacks.

### The Empirical Literature on Interracial Killings: Heterogeneity and Social Interaction

Almost all of the small number of empirical studies rely on Blau's (1977) heterogeneity theory (Sampson 1984, 1986; South and Messner 1986; O'Brien 1987; South and Felson 1990; Messner and Golden 1992; Messner and South 1992) because "with few exceptions, applications of Blau's theory to patterns of interracial violent crime have supported its propositions" (South and Felson 1990, p. 75). Blau argues that the proportion of group members is critical because these distributions determine the likelihood of social interaction between groups. Researchers therefore base predictions about levels of interracial crime on racial proportions. In cities with few blacks, the opportunities blacks have to interact with other blacks will be limited, but their contacts with whites should be enhanced.

These considerations help explain interracial homicide rates. If the percentage of blacks in a city is low, the reduced probability of black interactions with other blacks should produce increased black contacts with whites. This logic leads to an expectation that cities with few blacks will have higher rates of black killings of whites. Because interactions between potential black offenders and their white victims should be reduced in cities with higher percentages of blacks, the relationship between the percentage of blacks and the rate of black killings of whites should be negative.

The same logic suggests opposite relationships between the percentage of blacks and the rate of white killings of blacks. In cities with few blacks, potential white offenders will face limited opportunities to interact with blacks, so communities with larger percentages of blacks should have more killings of blacks by whites. In contrast to the expected negative relationship between the percentage of blacks and the rate of black killings of whites, the association between the percentage of blacks and the rate of white killings of blacks should be positive.

Residential segregation is another factor that must be included in any multivariate study of interracial homicides, because the rate of interracial killings depends on social contacts between the races. In cities where blacks and whites live apart, less interpersonal interaction can be expected. Social contact explanations suggest that residential segregation should be inversely related to rates of black killings of whites and white killings of blacks.

Interracial interactions are critical because they limit the possible

amount of interracial crime, but we wish to see if interracial economic and political conflicts explain interracial killings after social contact measures have been controlled. We begin our discussion of conflict accounts by presenting hypotheses about economic competition.

Conflict Inducements to Interracial Crime

*Economic competition.*—Competitive explanations for intergroup violence focus on the behavior that occurs when the gains of one group reduce the rewards of another (Hawley 1950). In order for groups to be in competition, they must interact. Groups that exist in separate spheres do not compete. Competition occurs only when groups occupy the same realized niche—a state called "niche overlap." When groups compete for the same niche, inevitably not everyone will find a place. Displacement occurs (Barth 1969), and this exclusion may produce hostility and violence. Studies of riots and lynchings suggest that competition for jobs leads to greater interracial violence (Beck and Tolnay 1990; Olzak 1990; Olzak 1992; Olzak et al. 1996), but we have not found studies that use such explanations to account for interracial killings.

The United States has a history of intense black subordination. Until recently these arrangements insured that the races occupied separate economic niches. Enhanced enforcement of stronger civil rights laws plus reduced racial gaps in education and skills recently have led to enhanced competition for the same jobs (Farley 1984; Donohue and Heckman 1991). With the number of positions held constant, as one racial group advances economically and obtains more jobs, employment opportunities for the other group *must* diminish. This zero-sum relationship means that as labor market competition between the races increases, the gap between large black and modest white unemployment rates will shrink, *because whites are losing jobs to blacks.*

In the recent past, black unemployment rates have been much higher than white rates partly because blacks have been confined to the bottom sector of dual labor markets (Dickens and Lang 1985, 1987), where crowding was intense and the resulting oversupplies led to black joblessness. As these racial barriers break down and blacks and whites begin to compete for the same positions, gaps between black and white unemployment rates should decrease because blacks are taking jobs from whites. But in cities where labor market segmentation on the basis of race persists, black unemployment rates should continue to be much greater than white unemployment rates. With other conditions held constant, where blacks cannot compete for many jobs, oversupplies of black labor will lead to enhanced black unemployment, but the same labor market protections will artificially increase demand for white workers and reduce white unem-

ployment. Yet in cities where white and black unemployment rates have become more equal, greater interracial competition will be present.

Following Olzak's (1992) logic, we use the black unemployment rate divided by the white unemployment rate to measure economic rivalries between blacks and whites. Black unemployment rates tend to be about twice as large as white rates. Cities with unemployment ratios that approach unity therefore should have the most economic competition between the races, but large scores on this ratio indicate less interracial competition. Because this measure is reversed so that low scores indicate enhanced competition, the racial unemployment ratio should be inversely related to interracial homicides if economic competition between blacks and whites leads to interracial killings. The greatest disparities in racial unemployment rates appear in those southern cities where black subordination is greatest and economic competition between the races is least likely, so the ratio of racial unemployment rates appears to be a valid measure of economic competition.[2]

The total unemployment rate should be controlled as well. Racial competition for jobs should be intensified in cities where total unemployment is high, but labor market competition will be reduced in cities with comparatively modest unemployment rates. According to Cohen and Felson (1979), unemployment has offsetting effects that should reduce this variable's association with crime. Yet the inclusion of the total unemployment rates in the models will increase the accuracy of the estimates of the relationship between interracial economic competition and interracial killings.

*Political competition.*—According to William J. Wilson, "racial strife has shifted away from the economic sector. The traditional racial struggles for power and privilege are now concentrated in the socio-political order. . . . The issues now have more to do with racial control over . . . *municipal* political systems than with the control of jobs" (1978, p. 121; emphasis added). If Wilson is right, the outcome of well-publicized political contests for the most important city office may alter racial tensions and influence interracial violence.

Many authors suggest that a significant part of black criminal violence

[2] The five cities with the highest ratio of black to white unemployment rates in 1980 are extremely traditional southern communities where it is probable that racially segmented labor markets have not been eroded as much as they have in other areas. These cities and their scores are Jackson, Miss., 3.88; Newport News, Va., 3.84; Mobile, Ala., 3.81; Macon, Ga., 3.63; and Montgomery, Ala., 3.47. The mean for this variable across all cities is only 2 19. Because these cites are among those with the most unaltered race relations, such scores indicate that the racial unemployment ratio appears to be greatest in areas where the economic subordination of blacks has changed least and black-white competition for the same jobs continues to be unusual.

directed against whites is a black reaction to their powerless status (Wilbanks 1985; LaFree 1982; Cleaver 1968). Perhaps where blacks gain political power, these conflicts are diminished. Clark (1994, p. 33) claims that the pattern of interracial conflict "shifts in U.S. cities with black mayors whose presence dramatically signals that there is no longer a white political elite." Although these political victories may be largely symbolic, well-publicized black successes in local elections evidently lead to reduced black feelings of powerlessness and perhaps fewer lethal conflicts with whites. Bobo and Gilliam (1990) find that blacks have feelings of heightened political efficacy and greater political influence if they live in areas with a black mayor. Lane (1997) claims that when blacks lost political influence, their homicide rates increased. Note as well that the number of interracial killings is not large. Hence, even small differences in the number of interracial killings partly attributable to black political victories in local elections should be enough to produce significant coefficients.

Equivalent reasoning suggests that a positive change in the political fortunes of blacks may increase white resentment and violence directed against blacks. Bobo and Hutchings (1996) report that whites are more likely to see blacks as political rather than economic rivals, while Bobo and Gilliam (1990) find that whites have diminished feelings of political efficacy in areas with a black mayor. Control over an apparently important political office by a minority can be viewed as a significant threat to the ascendant position of whites. If the primary political office in a city shifts to a minority, we can expect increased resentment by some members of the dominant group. In cities where blacks have gained sufficient political power to elect a black mayor, it is unlikely that those whites who are prone to interracial violence will be pleased. Yet control over such an apparently important political office may help curb violent black responses to their powerlessness.

These hypotheses about interracial political competition suggest relationships with opposite signs. In cities where the mayor is black, we should find higher rates of white killings of blacks because some whites should see the election of a black mayor as a reduction in the superordinate status of whites. Because the election of an African-American helps to reduce black feelings of powerlessness, we can expect an opposite relationship between the presence of a black mayor and black killings of whites. The latter argument suggests that the election of a black mayor will lead to decreased black killings of whites.

This racial threat approach also implies that in cities where there has been a recent growth in the percentage of blacks and a resulting increase in their political influence, blacks will feel empowered, so black killings of whites should be diminished. If this alternative political competition hypothesis is correct, change in the percentage of blacks in a community

will be negatively related to black killings of whites. An equivalent threat hypothesis predicts an opposite relationship when white killing rates of blacks are at issue. Blumer (1958) and Blalock (1967) suggest and Fosset and Kiecolt (1989) find that larger black populations threaten ascendant whites who may respond to these threats with violence. This group threat hypothesis suggests that the relationship between a recent increase in the percentage of blacks in a city and white killings of blacks should be positive.

### Police Size and Interracial Homicides

A larger police force may reduce some interracial killings more than others. Whites often have more political influence than blacks. As departments become larger, it is reasonable to expect that police resources may be differentially allocated to white neighborhoods. The logic of police work also may produce differential reductions in these killings. The police are best at controlling the streets largely because they cannot see through walls. Departments therefore have limited influence on crimes that occur within buildings (Rubenstein 1973; Stinchcombe 1963).

According to Parker and Smith (1979) and Parker (1989), in contrast to primary homicides, *nonprimary* homicides, or killings where the killer and the victim are not relatives, friends, or acquaintances, should be more responsive to formal social control. The data on the circumstances that contribute to interracial homicides in the supplemental homicide reports show that higher percentages of white killings of blacks occur during brawls or arguments, but most black killings of whites occur during felonies. Because brawls and arguments are more likely to occur within buildings (Stinchcombe 1963), the police will not be as effective in controlling these interracial crimes, so police per capita should not explain white killings of blacks.

For both reasons, it is reasonable to suspect that a larger police force can control felonies where the victim is white but the offenders are black. Even if the police are not more responsive to white demands for protection, stronger departments may disproportionately benefit whites because most black killings of whites occur during felonies rather than during interpersonal disputes. We therefore expect that larger police departments should reduce the black killings of whites, but police per capita may not be associated with white killings of blacks.

### Alternative Explanations

Although past researchers have neglected this control, any study that assesses the determinants of interracial killings should hold the total

murder rate constant. Interracial homicides are especially likely in cities with higher total homicide rates, but we want to isolate the unique determinants of interracial killings. We include total murder rates to control for a culture of violence and other explanations for noninterracial killings. This tactic reduces the possibility that explanatory variables will be significant because they explain total murders rather than interracial killings. Holding the murder rate constant also lets us focus on hypotheses about interracial killings and lets us neglect less relevant explanations from the vast literature on the more conventional intraracial homicides.

Interpersonal contacts will be both frequent and intense when many people live in a limited geographic area, so we hold constant another measure of the probability of interracial interaction by including population density in the equations. It is widely acknowledged that crime in populous cities is especially difficult to control. Many researchers have found that population is closely related to crime rates (Price 1966; Blau and Blau 1982; Jackson 1984). Higher interracial killing rates therefore can be expected in larger cities.

Because Sampson (1987) and others find that cities with higher percentages of single-parent families have more violent crimes, the percentage of female-headed families is entered in the Tobit models. When we analyze the rate of black killings of whites, we use the percentage of black families headed by a female. When the determinants of white killings of blacks are at issue, we instead use the percentage of white families headed by a female. Poverty may lead to interracial homicides, so we include the percentage of white or black families below the poverty line in appropriate analyses. Because regional cultural (and other) differences may influence interracial killings, we include dummies for three (census-defined) regions in some analyses.

This study analyzes relationships between latent racial conflicts and interracial homicide rates. We use the percentage of blacks in a city, residential segregation, and density to hold constant interracial interactions, and we look at the effects of economic competition between the races. Because political rivalries should matter, we see if the race of the mayor explains interracial killings. Change in the percentage of blacks may produce differing interracial killing rates, and we find out if larger departments reduce these rates. We also hold constant the total murder rate, density, city size, race-specific poverty rates, and race-specific percentages of families headed by a female. Finally, we control for cultural differences and other regional variations by including regional dummies in some equations. Table 1 summarizes the expected signs of the relationships.

TABLE 1

PREDICTED SIGNS FOR EACH TOBIT ANALYSIS

| Variable | Predicted Sign in Analyses of White Killings of Blacks* | Predicted Sign in Analyses of Black Killings of Whites* |
|---|---|---|
| ln %black ................................................................. | + | − |
| % change in %black ....... . .... ............................ | + | − |
| ln black/white %unemployed ............................. .. | − | − |
| %unemployed ....... . ... .... ...... . .. .     . ... . | + | + |
| ln population ......................................................... | + | + |
| ln murder rate ............... ..... ................................ | + | + |
| Police per capita ...... ................ ..... .. ............. .... . | 0 | − |
| Black mayor ........................... ........................... | + | − |
| √density .. ... ..... .......... . .. .. .... ... .. ................... . | + | + |
| % black families—female headed .. ... . .. ... | N.A | + |
| % poor black families ..... ........ .. ........ ......... . | N.A. | + |
| % white families—female headed ....................... | + | N.A. |
| % poor white families ... ...........-................ ......... .. | + | N.A. |
| Segregation .. ... ..     . .... .-...... . .... . . .. ... . | − | − |
| South . . . ... ... . . ... ... .............. ... ............. ..... | No prediction | No prediction |
| Northeast ................ .............-............................ .. | No prediction | No prediction |
| North central ............ ....... .......-................ . .. ... | No prediction | No prediction |

* N A means the variable in question will not be used in that analysis.

## METHODS

### The Data, Sample, and Estimation

The city-level data on interracial killings come from supplementary homicide reports filed by the police to the FBI (see Parker [1989] for strong arguments that cities are better units than metropolitan statistical areas for studying homicides). Because the unique conditions that lead to the use of lethal force by law enforcement officers are not at issue, killings by police officers have been removed. This information is likely to be accurate because homicides are difficult to conceal. We begin with the 170 cities with a population greater than 100,000 in 1980. Analyses are computed on 165 cities because residential segregation is not available for five cities.

   Interracial killings are infrequent events. To handle this problem, we compute the rates of black killings of whites and white killings of blacks in a city over a seven-year period starting in 1980 and ending at 1986. Such yearly aggregations have been used by others (Messner and Golden [1992] use five years). After averaging across these seven years, 32 cities (19.4%) still had no white killings of blacks, and 15 cities (9.1%) had no

black killings of whites.[3] Because OLS gives inconsistent estimates when the dependent variable is censored in this way, we estimate the models with Tobit.

Tobit uses two formulas for predicting the dependent variable—one for cases at the limit value (zero in this study) and another for cases above the limit—so it jointly estimates two effects (Roncek 1992, p. 503). In Tobit, probit estimates of the probability that a case will have a nonlimit value for cases at the limit on the dependent variable are combined with estimates of the effects of the explanatory variables on values of the dependent variable for cases with a nonlimit value on the dependent variable. We capitalize on the properties of the Tobit model because it is appropriate when the dependent variables are rates and these rates contain many zeros (Roncek 1992; Greene 1993; Long 1997).[4]

We analyze offender rates. We therefore compute the rate of black killings of whites by dividing the number of black killings of whites by the number of blacks in a city and multiplying by 100,000, while the rate of white killings of blacks is computed in the same fashion. To correct skewed distributions, the dependent variables are in natural log form (all logs are to the base $e$), but we follow convention and add one to avoid attempts to compute the log of zero.

Cities with extremely small black populations can have extremely high black killing rates of whites. If a city has a small black population, enormous rates can result from just one or two incidents over the seven-year period. To remove this distortion, we weight the Tobit models explaining black killings of whites by the number of blacks in a city. This procedure substantially reduces the effects of cities with few blacks while appropriately increasing the influence of cities with many blacks. To make the results comparable and to eliminate heteroscedasticity in the analyses of both dependent variables, we weight the equations that explain white killing rates of blacks by the number of whites in a city.[5] Reestimating the

[3] We deliberately do *not* average data for the explanatory variables over multiple years because some important independent variables (residential segregation for example) are available only in census years. Averaging an independent variable over many years almost always increases that variable's explanatory power. Because such means cannot be computed for each independent variable, they *must not* be used for any.

[4] For additional discussion and references to the many studies that have used Tobit, see Greene (1993) or any other standard econometrics text. For precedent in the use of this procedure by sociologists, see Walton and Ragin (1990), Roncek and Maier (1990), and Mosher and Hagan (1994). We use the Tobit routine in STATA, ver. 5.0.

[5] We follow the procedure recommended by STATA and use what these statisticians call analytic weights to correct for differing racial proportions. The total number of white killings of blacks for the seven-year period ranges from zero to 270 with a mean

standard errors with White's (1980) correction for heteroscedasticity suggests that this problem is not present.

## Measurement of Explanatory Variables

Economic competition is operationalized by the natural log of the ratio of the black to white unemployment rate. We correct skewness in the percentage of blacks with the natural log transformation. The percentage change in percentage of blacks is measured by using census data from 1970 and 1980, while city size is assessed with the log of population. We use a dummy scored "1" for the presence of black mayors.[6] To control for cultural differences and other regional spillovers, we sometimes include three regional dummies based on census definitions. Residential segregation is measured with a dissimilarity index computed by Taeuber and Monfort on block-level data in 1980 (see Messner and Golden [1992] for additional discussion).

Density is measured with residents per square mile in square-root form to correct for a mild skew. Police size is assessed with the police employees per 100,000 residents. We use all employees because departments substitute unsworn personnel for officers, but this choice does not matter. Murders are measured with the logged murder rate. Because we analyze offense rates, we use the percentage of black families with a female head and the percentage of black families below the poverty line when black

---

of 8.41. New York has the greatest number of these killings; Los Angeles is second with 121, and Chicago is third with 84 The number of black killings of whites ranges from zero to 382 with a mean of 15.73. Again, New York has the largest number; Los Angeles is second with 195, but Houston is third with 173 (table 2 gives the means and SDs of the relevant *rates*). The large number of interracial killings in these and other cities rules out procedures like Poisson regression that are inappropriate for estimating counts beyond about 50 (see Maddala [1977] for reasons why such count estimators are unsuitable when dependent variable scores reach maximums significantly greater than 50).

[6] Of the cities in the analysis period, 11 had black mayors. Cities with black mayors tended to be larger and more dense than cities with white mayors. The percentage of blacks in cities with a black mayor was greater than it was in the remaining cities, but the degree of residential segregation was almost identical. The characteristics of white or black residents in cities with a black or white mayor are similar. For example, black median family incomes were $14,102 in cities with a white mayor, while this median was $13,401 in cities with a black mayor White median family incomes were $21,272 in cities with a white mayor, but they were $21,824 in cities with black mayors. Although there are some differences between these cities, these differences are held constant in the Tobit models. Note as well that the black mayor dummy had extremely weak correlations with the other explanatory variables in the models, so the political competition findings probably are not produced by collinearity

killings of whites is analyzed. When we analyze white killings of blacks, we include the percentage of white families headed by a female and the white family poverty rate.

## Model Specifications

When we predict the rate of white killings of blacks, the coefficients on three explanatory variables should be negative. The black-white unemployment ratio and residential segregation should be inversely related to these killings, but the remaining Tobit coefficients should take positive signs, and the police variable should not be significant. One of the more general specifications of the weighted (by the number of whites) Tobit model that explains rates of white killings of blacks is

$$\text{WHTKILLB}_i = b_0 + b_1\text{PERBLK}_i + b_2\text{BLK/WHTUN}_i + b_3\text{UNEM}_i$$
$$+ \, b_4\text{MURDRT}_i + b_5\text{POLSIZ}_i + b_6\text{BLACKMAYOR}_i$$
$$+ \, b_7\text{POP}_i + b_8\text{SEG}_i + b_9\text{CHPERBLK}_i$$
$$+ \, b_{10}\text{DENSITY}_i + b_{11}\text{WHTFEMLHD}_i$$
$$+ \, b_{12}\text{SOUTH}_i + b_{13}\text{NEAST}_i + b_{14}\text{NCENT}_i,$$

where WHTKILLB is the natural log of white killings of blacks plus one per 100,000 white residents, PERBLK is the natural log of the percentage of blacks, BLK/WHTUN is the natural log of the ratio of the black to white unemployment rate, UNEM is the total unemployment rate, MURDRT is the natural log of total murders per 100,000 population, POLSIZ is the per capita number of police, BLACKMAYOR is a dummy coded "1" if the mayor is black, POP is the natural log of city population, SEG is the residential segregation index, CHPERBLK is percentage change in the percentage of blacks, DENSITY is the square root of the number of persons per square mile, WHTFEMLHD is the percentage of white families headed by a female, SOUTH is a dummy for the South, NEAST is a dummy for the Northeast, and NCENT is a dummy for the north-central region. In an additional model, we enter the percentage of white families below the poverty line.

When we analyze the rate of black killings of whites, six explanatory variables should take negative signs. The percentage of blacks, the ratio of black to white unemployment, police per capita, the presence of a black mayor, residential segregation, and change in percentage of blacks should be inversely related to black killings of whites, but the coefficients on the remaining explanatory variables should be positive. One of the more general specifications of the weighted (by the number of blacks) Tobit model

TABLE 2

MEANS AND STANDARD DEVIATIONS OF ALL VARIABLES

| Variable | Mean | SD |
|---|---|---|
| ln black killings of whites rate................... ......... ... .. | 3.0584 | 1.2907 |
| ln white killings of blacks rate . .. .. .................... | 1.0210 | .68302 |
| ln %black . ................................----------- . ........................... | 2.3648 | 1.3245 |
| % change in %black ........................ .. .. ..... ....... | 64 201 | 174.70 |
| ln black/white %unemployed ..................... .. ... . .. | .74914 | 28223 |
| % unemployed ..................----- . ... .. .. ............. | 6.8283 | 2.8076 |
| ln population . ..................... ............... .. . .. ............'.. | 12.303 | .75248 |
| ln murder rate ... ....... ........... ............................ ...... | 2.4512 | .83964 |
| Police per capita · ..... .............................. ...... ......... | 257.86 | 95.939 |
| Black mayor ..........................-. .. ... .. .. .................. ... . | .06061 | 23933 |
| √density .. ...... . ...................... ... .. .. ... ................ ... | 19.757 | 7.0762 |
| % black families—female headed .. ... .. .............. | 27.716 | 6.8514 |
| % poor black families ...... ..................... .... . ... ...... . | 23.322 | 7.5475 |
| % white families—female headed ................. .. ... | 8.1615 | 1.5963 |
| % poor white families ................. .. . .. .................. | 6.9012 | 2 6158 |
| Segregation ...... .. .. ... ... -- .. .. ................. .. .. .. ....... | 72 530 | 11.195 |
| South .. ....................... ... ... .. .................... .... . . ...... | .35152 | .47890 · |
| Northeast . ......................... .. . .. .. ...................... .... | .10999 | .31270 |
| North central . .. ................. ..... .. ... .. .. ............... ..... | .23030 | .42231 |

NOTE.—$N = 165$

that predicts rates of black killings of whites therefore is

$$\text{BLKKILWHT}_i = b_0 + b_1\text{PERBLK}_i + b_2\text{BLK/WHTUN}_i + b_3\text{UNEM}_i$$
$$+ b_4\text{MURDRT}_i + b_5\text{POLSIZ}_i + b_6\text{BLACKMAYOR}_i$$
$$+ b_7\text{POP}_i + b_8\text{SEG}_i + b_9\text{CHPERBLK}_i$$
$$+ b_{10}\text{DENSITY}_i + b_{11}\text{BLKFEMLHD}_i$$
$$+ b_{12}\text{SOUTH}_i + b_{13}\text{NEAST}_i + b_{14}\text{NCENT}_i,$$

where BLKKILWHT is the natural log of black killings of whites plus one per 100,000 black residents, BLKFEMLHD is the percentage of black families headed by a female, and all other variables are defined as above. In another model, we enter the percentage of black families below the poverty line.

RESULTS

Analyses of White Killings of Blacks

Table 2 shows means and standard deviations. Table 3 gives the results of the weighted Tobit models that explain the rate of white killings of

blacks. The first model is restricted. We then include additional explanatory variables in subsequent models by adding percentage change in the percentage of blacks, density, the white family poverty rate, and the percentage of white families headed by a female to the next equations. In the last model, we add the three regional dummies.

As one would predict from the past literature that focused on the probability of interracial interactions, the percentage of blacks is positively related to white killings of blacks. This finding makes sense because black victimization rates should be modest in cities with few blacks. The strong relationship between total murder rates and these interracial killings suggests that past researchers should have included the murder rates in their models.

Other regressors have negligible or inconsistent effects. Race-specific poverty rates, percentage female-headed white families, aggregate unemployment, population, density, and region do not matter. It is noteworthy that the relative size of the police departments is unrelated to white killings of blacks. Residential segregation and percentage change in the percentage of blacks also have negligible associations with the likelihood that whites will kill blacks.

Two theoretically important findings represent departures from past research. First, the results show that higher rates of white killings of blacks occur in cities where the black unemployment rate approaches the unemployment rate of whites. This persistent finding suggests that threats due to interracial economic rivalries lead to more white killings of blacks. Second, we find that cities with a black mayor have more white killings of blacks, so this rate is greater in cities where the most visible political office has shifted to an African-American. But we do not know if these results will hold when the opposite interracial homicide rate is analyzed.

Analyses of Black Killings of Whites

Table 4 shows the weighted Tobit models predicting the rate of black killings of whites. The first restricted model in table 4 is identical to model 1 in table 3, and these similarities are repeated. In table 4, we follow the pattern in table 3 and add percentage change in the percentage of blacks, density, the black family poverty rate, and percentage of black female-headed families to subsequent models. In the last model, we again include the three regional dummy variables.

In contrast to its effects on white killings of blacks, the percentage of blacks is inversely related to the rates of black killings of whites, but this negative association is expected because minorities tend to interact with members of their own race. A finding that cities with greater percentages of blacks have fewer black killings of whites supports past research about

TABLE 3

WEIGHTED TOBIT EQUATIONS ASSESSING THE DETERMINANTS OF WHITE KILLINGS OF BLACKS

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Intercept | -.7869 | -.6601 | -.3778 | -.5240 | -.4539 |
| | (-1.87) | (-1.52) | (-.80) | (-.90) | (-.76) |
| ln %black | .2788*** | .2702*** | .2628*** | .2625*** | .2577*** |
| | (5.16) | (4 96) | (4.84) | (4.83) | (4.31) |
| ln black/white %unemployed | -.4626** | -.5005** | -.4511** | -.4893** | -.5038** |
| | (-2.82) | (-2 99) | (-2.66) | (-2.74) | (-2 90) |
| %unemployed | -.0054 | -.0063 | -.0101 | -.0093 | -.0140 |
| | (-.40) | (-.47) | (-.75) | (-.65) | (-.93) |
| ln murder rate | .4365*** | .4328*** | .4533*** | .4686*** | .4455*** |
| | (5.90) | (5.85) | (6.05) | (6.05) | (5 50) |
| Police per capita | .0000 | .0001 | -.0002 | -.0002 | -.0000 |
| | (.11) | (.21) | (-.48) | (-.52) | (-.15) |
| Black mayor | .1816* | .1818* | .2177* | .2065* | .2233* |
| | (1.67) | (1.68) | (1.97) | (1.86) | (1.92) |
| ln population | .0502 | .0469 | -.0021 | .0062 | .0138 |
| | (1.64) | (1.53) | (-.54) | (.15) | (.34) |

| | | | | | |
|---|---|---|---|---|---|
| Residential segregation | −.0031 (−.84) | −.0033 (−.89) | −.0007 (−.18) | −.0010 (−.25) | −.0049 (−1.10) |
| % change in %black | ... | −.0003 (−1.10) | −.0003 (−1.20) | −.0004 (−1.33) | −.0004 (−1.48) |
| √density | ... | ... | .0070 (1.41) | .0081 (1.56) | .0124* (2.15) |
| % white families—female headed | ... | ... | . | .0216 (.67) | .0136 (.48) |
| % poor white families | . | ... | | −.0144 (−.80) | .. |
| South | ... | ... | | ... | .1273 (1.19) |
| Northeast | ... | ... | | ... | −.1324 (−.96) |
| North central | ... | ... | | .. | .1434 (1.27) |
| Log likelihood | −78.06 | −77.42 | −76.43 | −76.07 | −73.54 |
| Chi-square test | 201.78*** | 203.05*** | 205.03*** | 205.76*** | 210.81*** |
| Pseudo R² | .564 | .567 | .573 | .575 | .589 |

NOTE.—*t*-values are given in parentheses.
* $P \le .05$
** $P \le .01$.
*** $P \le .001$

173

## TABLE 4

### WEIGHTED TOBIT EQUATIONS ASSESSING THE DETERMINANTS OF BLACK KILLINGS OF WHITES

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Intercept | 4.3151*** | 4.5277*** | 4.4817*** | 3.7550*** | 3.5731*** |
| | (10.54) | (10.94) | (9 89) | (7.22) | (6.39) |
| ln %black | -.7445*** | -.7612*** | -.7567*** | -.7504*** | -.7084*** |
| | (-11.37) | (-11.52) | (-12.46) | (-10.75) | (-8.63) |
| ln black/white %unemployed | -.3454* | -.3385* | -.3572* | -.3705* | -.4478** |
| | (-2.10) | (-1.98) | (-2.07) | (-2.15) | (-2.59) |
| %unemployed | -.0137 | -.0106 | -.0103 | -.0158 | -.0107 |
| | (-1.33) | (-1.03) | (-.98) | (-1.52) | (-.86) |
| ln murder rate | .6505*** | .6470⁴** | .6438*** | .6850*** | .6407*** |
| | (10.56) | (10.17) | (10.03) | (9 93) | (9.00) |
| Police per capita | -.0012*** | -.0012*** | -.0011*** | -.0010** | -.0010** |
| | (-4.09) | (-4.05) | (-3.21) | (-2.80) | (-2.99) |
| Black mayor | -.2380** | -.2455*** | -.2569** | -.2976*** | -.3440*** |
| | (-2.77) | (-2.86) | (-2.86) | (-3.33) | (-3.87) |
| ln population | .0405 | .0373 | 0461 | .0743* | .0826* |
| | (1.50) | (1.38) | (1.18) | (1.87) | (2 07) |

| | | | | |
|---|---|---|---|---|
| Residential segregation | −.0088* | −.0101* | −.0107*** | −.0129** | −.0110* |
| | (−2.27) | (−2.37) | (−2 40) | (−2.73) | (−2.26) |
| % change in %black | . . | −.0014*** | −.0014**** | −.0017**** | −.0013*** |
| | | (−4.68) | (−4 60) | (−5.19) | (−4.23) |
| √density | ... | . | −.0016 | −.0075 | −.0057 |
| | | | (−.33) | (−1 38) | (−.95) |
| % black families—female headed | ... | . | ... | .0207** | .0207** |
| | | | | (2.39) | (2.41) |
| % poor black families | ... | . | | −.0019 | |
| | | | | (−.24) | |
| South | ... | . | ... | ... | −1253 |
| | | | | | (−1.02) |
| Northeast | ... | . | . | | −2406* |
| | | | | | (−1.90) |
| North central | . | . | . | . | −.2079* |
| | | | | | (−2.04) |
| Log likelihood | −72.56 | −70.96 | −70.91 | −68 23 | −66.89 |
| Chi-square test | 171.59**** | 174.80**** | 174.89**** | 180 25*** | 182 92*** |
| Pseudo $R^2$ | .542 | .552 | .552 | .569 | .578 |

NOTE.—$t$-values are given in parentheses.
* $P \leq .05$,
** $P \leq .01$.
*** $P \leq .001$

the determinants of interracial killings. The significant negative coefficient on residential segregation also is consistent with contact explanations. We again find that the total murder rate predicts interracial killings, but some control variables again have negligible effects. Total unemployment, density, and the black poverty rate do not explain black killings of whites, but populous cities may have more of these killings. This conclusion must remain tenuous because the population variable is not significant in the first models in table 4.

Economic and political rivalries continue to explain these interracial killings. Cities with greater economic competition between the races have more black killings of whites as well as higher rates of white killings of blacks. In accord with predictions, the relationship between the presence of a black mayor and the rate of black killings of whites is opposite from the relationship between black mayors and white killings of blacks. The coefficients in table 3 show that cities with a black mayor have higher rates of white killings of blacks, yet these same cities have fewer black killings of whites. Such results imply that the election of a black mayor reduces black lethal violence directed against whites. This political effect is corroborated by a second finding. Cities that have experienced higher growth rates in black population have reduced black killings of whites. Both results suggest that black killings of whites are less likely in cities where black political influence has increased.

The coefficients on police strength are equally intriguing. The results in table 3 show that the per capita number of police employees is unrelated to white killings of blacks, but the significant negative coefficients on this variable in all models in table 4 suggest that cities with stronger departments have diminished black killings of whites.

Additional Tests

If the presence of a black mayor alters the interracial killing rates because black political victories reduce black feelings of powerlessness (Bobo and Gilliam 1990), but they increase white resentment against blacks, we would not expect these political victories to explain black killings of blacks or white killings of whites. Interracial resentments should not influence within-race killings. Because interracial competition for jobs should enhance racial conflict and because it is difficult to see how such conflicts would lead to within-race killings, the relationships between economic competition and white killings of whites or black killings of blacks should be weaker than the relationships between economic competition and white killings of blacks or black killings of whites.

The supplemental Tobit analyses reported in the appendix support these expectations. They suggest that neither the presence of a black

mayor nor economic competition between blacks and whites leads to increased white killings of whites or black killings of blacks. The robust relationships between political and economic competition and interracial killings reported in tables 3 and 4 combined with the negligible associations between either political or economic competition and within-race killing rates increase the plausibility of our competition explanations for interracial killings.

Other alternative explanations probably can be eliminated. Balkwell (1990) finds that racial inequality explains homicides. Perhaps the significant coefficients on the unemployment ratio are due to racial differences in economic resources. To assess this alternative hypothesis, we include the ratio of black to white median or mean family incomes in unreported Tobit models. The coefficients on this ratio are negligible, however, and the inclusion of this variable does not alter the significance tests on any of the coefficients reported in tables 3 and 4.

Although the racial unemployment *ratio* explains interracial killings, the total unemployment *rate* does not, perhaps because the total unemployment rate has offsetting effects. According to Cohen and Felson (1979), unemployment produces crime by increasing economic need, but it also keeps people near their homes so interracial contacts will be reduced. Yet as one would expect from an argument that labor market competition between the races becomes more intense in cities with substantial total unemployment rates, we find that including the total unemployment rate strengthens the coefficients on the unemployment ratios, and this finding is not due to collinearity.[7]

The results persist if other variables are included in unreported equations. Income inequality measured by Gini, %change in the unemployment rate, and the percentage of dwelling units with more than 1.01 persons per room have negligible effects.[8] If we use either %black or %whites ages 15–24, these variables do not explain the interracial killing rates, and the theoretical implications are left unaltered by these controls. All other

[7] The correlation between the unemployment ratio and total unemployment weighted by the number of blacks is −.29; this correlation is −.02 when the number of whites is used as a weight, while the unweighted correlation is −.05. Attempts to use the ratio of black to white employment (rather than unemployment) in log or unlogged form met with no success.

[8] Inequality measures such as Gini computed on city incomes should have offsetting effects that produce insignificant coefficients. Because blacks tend to be much closer to the bottom of the income distribution than whites, heightened inequality could generate enhanced interracial conflicts for resources. But wider gaps in racial distribution of economic resources should increase social distance between the races and decrease interracial social contacts, so opportunities for interracial homicides would be reduced in unequal cities. These countervailing effects should cancel, so the negligible associations we find between Gini and interracial killings are not surprising.

significant findings also remain if the black mayor variable is dropped from the models, so the additional findings do not depend on the inclusion of this controversial explanatory variable.

Model 5 in both tables shows that the results persist after three regional dummy variables are entered. Increasing the specificity of these regional measures by using eight dummies to represent the nine census subregions has the same effect in unreported analyses. *These findings make it extremely difficult to believe that the results depend on region in any way.*

Corrections for heteroscedasticity using White's (1980) method do not alter the theoretical implications of the results. Finally, the stability of the estimates after substantial changes in the explanatory variables included in the models suggests that collinearity is not a problem. *All* VIF scores are below the threshold of four that the most conservative statisticians advocate as an indicator of collinearity.[9]

We find that larger percentages of blacks, enhanced economic competition, and the presence of a black mayor lead to more white killings of blacks. The results show that the percentage of blacks, residential segregation, the size of the police force, the presence of a black mayor, and increases in percentage of blacks reduce black killings of whites, but cities with enhanced interracial economic competition also have more black killings of whites. Perhaps racial conflicts are not influential enough to overcome the limited amount of contact between blacks and whites and make interracial killings as widespread as intraracial killings. The results nevertheless support hypotheses that economic and political conflicts between blacks and whites lead to both white killings of blacks and black killings of whites after contact measures are held constant.

DISCUSSION AND IMPLICATIONS

These findings depart from past studies, but the differences probably appear because we use better methods and new explanatory variables. First, we study disaggregated interracial killings. This lets the results differ for the opposite interracial killing rates. Second, we use Tobit to correct for censoring. Third, we weight the analyses to eliminate heteroscedasticity and reduce the influence of small cities with few blacks. Fourth, we con-

---

[9] To see if influential cases are distorting the results, we use the robust *regression* routine in STATA (according to Long [1997], an equivalent procedure does not exist for Tobit). To avoid conclusions that rest on a few cases, this procedure weights influential cases less heavily. When we compare the robust regression estimates to those produced by OLS, the same variables are significant in each analysis, so this indirect (it is not conducted with Tobit) procedure suggests that the Tobit findings probably do not rest on a few influential cases.

trol for total murder rates and thereby reduce the likelihood of spurious relationships. Finally, the large number of regressors and the persistent results despite severe changes in the models increases the credibility of the findings.[10] Because the black mayor variable remains significant with the expected opposite signs after many explanations including region have been held constant, these relationships probably are not spurious.

## Results Consistent with Heterogeneity Theory

As expected, the likelihood of interracial contact as measured by the percentage of black residents is a consistent predictor of the probability of interracial killings with signs in the anticipated direction. In cities with relatively few blacks, the reduced probability of black interactions with other blacks should lead to more black interactions with whites and higher rates of black killings of whites. This reasoning leads to an expectation that the association between the percentage of blacks and black killings of whites will be negative, and that is what we find.

The same logic leads to a prediction that the relationship between the percentage of blacks and the rate of white killings of blacks should be positive. Potential white offenders will face limited opportunities to interact with black victims where there are few blacks. Hence, cities with larger percentages of blacks should have higher white killings of black rates, and the results consistently fit with this expectation as well.

These findings support the emphasis on explanations derived from Blau's (1977) heterogeneity theory in the literature on interracial killings, but most of these investigators ignored 'conflict explanations. While these interracial contact indicators are important probably because they limit the amount of interracial killings, that does not mean that such contacts are the only explanation for these lethal events. When we look at the effects of economic and political rivalries between blacks and whites and the dissimilar consequences of police department size, we find that hypotheses derived from conflict theory consistently explain variation in interracial killing rates after heterogeneity explanations have been taken into account.

---

[10] Johnston's claim about the number of regressors in models is worth quoting. He says, "It is more serious to omit relevant variables than to include irrelevant variables since in the former case the coefficients will be biased, the disturbance variance overestimated, and conventional inference procedures rendered invalid, while in the latter case the coefficients will be unbiased, the disturbance variance properly estimated, and the inference procedures properly estimated. This constitutes a fairly strong case for including rather than excluding relevant variables in equations. There is, however, a qualification. Adding extra variables, be they relevant or irrelevant, will lower the precision of estimation of the relevant coefficients" (1984, p. 262). So inclusive specifications typically give more conservative significance tests.

## Results Consistent with Findings about Other Forms of Interracial Violence

The results invariably suggest that competition for jobs between the most subordinate and the most dominant race lead to more interracial killings as well. The black-white unemployment ratio has the predicted negative relationships with both white killings of blacks and black killings of whites. Such inverse relationships are what one would expect if economic competition between the races leads to interracial killings. These consistent findings suggest that past researchers should not have overlooked the possibility that economic rivalries explain at least some interracial killings.

A conclusion that economic competition between blacks and whites leads to interracial homicides does not contradict research findings about other forms of interracial violence in the past. Researchers who use historical evidence claim that economic competition in the southern United States led to increased violence that disproportionately victimized African-Americans (Bonacich 1972, 1975; Williamson 1984).

Statistical studies support this conclusion as well. Beck and Tolnay (1990) find that lynchings were likely when economic competition between the races became more intense, while Olzak (1992) shows that interracial economic competition increased the probability of race riots. Our findings that economic competition between blacks and whites explains both kinds of interracial killings support these conclusions even though the prior findings are based on other kinds of interracial violence that occurred during other historical periods. Contrasts in the methods and data employed in these diverse studies make the equivalent results especially noteworthy.

## New Findings

The results show that cities with larger police departments have reduced rates of black killings of whites, but departmental size has no effects on white killings of blacks. Such findings represent a departure from the prior work on the deterrent effects of department size. Only a few studies have found the expected negative relationship between the per capita number of police officers and violent crime rates (see Sampson and Cohen [1988] for a discussion).

Our unusual findings probably emerge because the two types of interracial killings differ. Black killings of whites are most likely to happen during a felony, but white killings of blacks generally occur during interpersonal disputes. Because black felons in white neighborhoods should be conspicuous and because the police rely on "that which is out of place" to make decisions about who should be watched or stopped, it is not surprising that cities with larger departments would have fewer black killings

of whites. It is equally plausible that department size would not be related to white killings of blacks because the ability of the police to control violence that arises from interpersonal disputes is more limited (Parker and Smith 1979).

Yet the finding about political rivalry is the most novel. With many factors held constant, the results invariably show that cities with black mayors have higher rates of white killings of blacks but fewer black killings of whites. We also find that black killings of whites are reduced in cities where the black population has recently increased. Particularly because the presence of a black mayor is positively associated with black perceptions of political efficacy and influence (Bobo and Gilliam 1990), these findings do not contradict speculations that some black violence directed against whites is due to black powerlessness.

Parallel findings showing that white killings of blacks are more likely in cities with black mayors support the racial threat literature by suggesting that some whites respond to the menace posed by increased black political influence with enhanced violence directed against this previously more subordinate minority. Yet these interpretations of the political findings must remain tentative because we do not have measures of perceptions.

Such opposite results that fit expectations derived from conflict theory are intriguing because they support speculative claims that interracial criminal violence is influenced by contests for power and by the degree of subordination of one race by another. Our findings suggest that political rivalries spill over to influence interracial killings, but this form of quasi-political criminal violence has been overlooked in the empirical literature on interracial killings. The persistence of the opposite-signed coefficients on the black mayor variable despite drastic alterations to the equations and despite the large number of diverse effects held constant makes it difficult to believe that these relationships are spurious.

It is important to remember that the reported estimates represent independent effects. While cities with larger black populations are more likely to have a black mayor, the coefficients are adjusted for this and other associations between explanatory variables. Another possible criticism runs that black mayors would not be pleased by increased white killings of blacks during their tenure Yet we find that the unique effects of the presence of a black mayor are positive. Perhaps the causes of white killings of blacks during the tenure of a black mayor are powerful enough to overcome the public policies black mayors may institute to reduce white killings of blacks.

Such partially offsetting effects do not seem likely for several reasons. The literature suggests that few (or perhaps no) anticrime policies put forward by mayors and other political officials have been even modestly

effective. Because readily available statistics about the recent incidence of interracial killings do not exist, we wonder if mayors would have any reliable information about such conditional shifts in the interracial homicide rates. While it is plausible that city officials have some knowledge about recent changes in the aggregate murder rates in their jurisdictions, it is unlikely that they would have accurate information about the far less visible shifts in the separate interracial homicide rates.[11] Nevertheless, the novelty of this intriguing black mayor finding suggests that it should be subjected to further investigation.

### Wider Implications

Although a few researchers have tried to use economic inequality to explain interracial killings, hypotheses that stress political and economic conflicts between the races have not been subjected to as much empirical scrutiny. This omission is unfortunate because it is difficult to believe that the antagonisms stemming from economic and political struggles between dominant and subordinate racial groups do not account for at least some of the lethal violence that occurs between the races. Because we use appropriate methods and include new measures of these effects, we find evidence that struggles for political power and jobs should not be overlooked when researchers try to isolate the factors that produce violent interracial crimes.

Empirical efforts to test the conflict approach in criminology focus on the behavior of social control agencies, but the incidence of crime has been

---

[11] The absence of accurate information about these race-specific killing rates also makes it hard to believe that simultaneity is biasing the coefficients. While the media may publicize a few unusually vivid interracial killings, it is implausible that reliable knowledge about the actual likelihood of these relatively infrequent events would be sufficiently widespread to systematically influence either racial migration or voting in local elections. Studies have shown that the attention the media gives to crime has little or no relationship with the amount of crime (Fishman 1980; Orcutt and Turner 1993). The only publicly available information about recent interracial killings is likely to create mistaken perceptions about these risks. Beckett (1997) uses time-series regressions and finds negligible associations between the far more visible FBI crime rates (that have not been separated by race of victims and offenders) and public concerns about crime. Because accurate information about recent interracial killing rates is impossible to obtain, it is not easy to see how these race-specific interracial homicide rates could lead to differential racial migration or shifts in white or black voting. But it is possible to see if this alternative explanation is distorting the results. Equations that include the appropriate 1980 interracial killing rates as lagged explanatory variables give theoretically identical results The independent variables that were significant in tables 2 and 3 continue to be significant after lagged dependent variables are used on the right-hand side of the equations. Such findings suggest that simultaneity is not problematic.

largely ignored. Conflict theorists have studied the determinants of the size or the resources of police departments (Jacobs 1979; Liska, Lawrence, and Benson 1981; Jackson 1989; Jacobs and Helms 1997), the use of deadly force by the police (Sherman and Langworthy 1979; Jacobs and O'Brien 1998), the determinants of arrest rates (Gove, Sullivan, and Wilson 1998), and fluctuations in imprisonments (Parker and Horwitz 1986; Chiricos and Delone 1992; Jacobs and Helms 1996). But empirical studies that use conflict explanations to successfully explain the incidence of crime are unusual.

This perspective is not easy to apply to most street crimes because the victims are so often from the underclass. Such victimization patterns present an obstacle to social scientists who see crime as inarticulate rebellion. Those who use violence to alter the existing social order should aim at the powerful rather than targeting people with the least influence. Because this study focuses on crimes where the offender and the victim come from subordinate and superordinate races and because we examine new explanations that involve political and economic divisions, we find evidence that these lethal crimes can be explained by interracial economic and political rivalries.

A paradox remains. Most of the research on state control agencies motivated by conflict theory finds that greater inequality in one form or another leads to enhanced control. Yet findings from research on racial competition in this study and by people like Olzak and her associates (1990, 1992, 1996) or Beck and Tolnay (1990) suggest that intergroup violence expands as competition between the races increases and blacks and whites become more equal. These diverse findings suggest that perhaps heightened inequality creates greater opportunities for elites to use the state to control unruly behavior that threatens their position. Yet reduced inequality between the nonelite members of dominant groups and groups who once were in extremely subordinate positions leads to enhanced intergroup violence.

The findings in this study can be briefly summarized: They show that Blau's heterogeneity approach that has been given so much attention in the literature certainly helps to explain the incidence of interracial crimes. Yet this perspective is not sufficient. First, it ignores competition between the races for scarce but valuable resources like jobs and political power. Second, this perspective overlooks social control agencies like the police who evidently can protect members of the more dominant race better than they can protect minorities. Our results suggest that stronger state agencies that specialize in social control disproportionately help members of the most influential race.

Racial conflict probably is not powerful enough to overcome infrequent social contacts between the races and make interracial killings as common

as within-race killings. Nevertheless, if we restrict our focus to just these unusual but theoretically intriguing events, in contrast to past studies, we find that the racial conflicts created by a history of economic and political subordination account for differences in the relative rates of white killings of blacks and black killings of whites. These findings suggest that researchers should pay greater attention to economic and political rivalries as explanations for violent crimes, particularly if offenders and their victims come from social groups that have not been treated equally.

APPENDIX A

Justification and Methods

To see if the expected negligible relationships between political or economic competition and within-race killings are present, we report supplemental Tobit analyses of the rates of black killings of blacks and white killings of whites. Comparisons of the determinants of intra- versus interracial killing rates are at issue, so we duplicate the specifications and the methods used in our models that explain interracial killing rates as much as possible. These equations should resemble those used in the analyses of interracial killings because the dependent variables measure homicides, yet within-race killings differ from interracial killings. The specifications used in these supplemental analyses of intraracial killings therefore should not be identical to those used in the reported models designed to explain interracial killings, but they ought to be similar.

The explanatory variables selected for the models developed to explain within-race homicides must have theoretically plausible relationships with such killings. Variables that are not used in the prior analyses of interracial killings because they are insignificant must explain at least one of the within-race killing rates to be included in these supplemental models. For this reason, we add a Gini index computed on household incomes to gauge the degree of economic inequality in these cities and crowding measured by the natural log of the percentage of dwelling units with more than 1.01 persons per room. (Recall that the inclusion of these variables in the models that explain between-race killings does not have any effects on the theoretical implications of the results). Because there is no reason to think that racial segregation would influence white killings of whites (and because segregation is insignificant in that equation), yet it is plausible that segregation would lead to increased black killings of blacks, we use residential segregation only in the analyses of black killings of blacks.

Within-race killings are far more common than between-race killings, so it is not necessary to average scores on the dependent variable over multiple years. We therefore use data on within-race homicides collected by the Supplemental Homicide Reports for 1982. To facilitate compari-

sons between the determinants of intraracial killings and interracial kill-
ings and to adjust for the modest censuring that is present in these two
within-race homicide rates, we use the same weighted Tobit procedure
that we use to estimate the determinants of interracial homicides. The
dependent variables in these supplemental analyses are the rates of black
killings of blacks and white killings of whites in natural log form with
one added to both rates to avoid attempts to compute the log of zero.

## Analyses

In the first two models in table A1, we report findings from the analyses
of white killings of whites. Model 1 presents a restricted analysis, but in
model 2 we add the three additional predictors that make this model simi-
lar but not identical to model 5 in table 3. The remaining two models show
the relationships between comparable explanatory variables and black
killings of blacks using race-specific explanatory variables computed with
blacks rather than whites. In model 3, we use a restricted specification
similar to that used in model 1 of table A1, but in model 4 we add the
remaining three variables that make this model similar to model 5 in
table 4.

   These results show that neither political nor economic competition ex-
plain within-race killing rates. The coefficients on the black mayor vari-
able are negligible in all four equations, while the coefficients on the racial
unemployment ratio that measures the degree of interracial competition
for jobs also never reach statistical significance. These expected negative
findings persist when we use alternative (but unreported) models. Finally,
it also is worth noting that police per capita does not explain either of the
within-race killing rates.

## Conclusions

If the presence of a black mayor leads to increased racial friction due to
enhanced white resentments against black political successes, but black
feelings of powerlessness are reduced when cities have a black mayor
(Bobo and Gilliam 1990), jurisdictions with a black mayor should have
more white killings blacks but fewer black killings of whites. Yet this
indicator of political conflict between the races should not explain black
killings of blacks or white killings of whites because it is difficult to see
how political conflicts between the races would lead directly to within-
race killings. The results from our Tobit analyses of each of the four differ-
ent race-specific killing rates are consistent with these expectations.

   If competition for jobs increases interracial conflict, these rivalries
should lead to more white killings of blacks and black killings of whites.
Yet they should not have equally strong relationships with white killings

## TABLE A1

### WEIGHTED TOBIT EQUATIONS ASSESSING THE DETERMINANTS OF WITHIN-RACE KILLINGS

| | WHITE KILLINGS OF WHITES | | BLACK KILLINGS OF BLACKS | |
|---|---|---|---|---|
| | Model 1 | Model 2 | Model 3 | Model 4 |
| Intercept | -2 1447* | -2.4355** | -4 1610*** | -4.1739*** |
| | (-2.55) | (-2.88) | (-4.01) | (-4.02) |
| %white or ln %black[a] | 0070* | .0042# | .4380*** | .4410*** |
| | (1.84) | (1 82) | (3 72) | (3.53) |
| ln black/white %unemployed | -.1553 | -.2048 | .0840 | -.0153 |
| | (-.96) | (-1.24) | (.34) | (-06) |
| %unemployed | -.0231 | -.0273 | .0754*** | .0672*** |
| | (-1.37) | (-1.61) | (3 94) | (3.36) |
| Black mayor | -.1086 | -.1326 | .1283 | .1250 |
| | (-86) | (-1.05) | (.81) | (.76) |
| ln population | .1469*** | .1479*** | -.1016* | -.1013* |
| | (3.89) | (3.90) | (-1.91) | (-1.86) |
| Income inequality | 2.4158 | 2.8380* | 8.8239*** | 9.1736*** |
| | (1 57) | (1.76) | (3 14) | (3.14) |
| √density | -.0167*** | -0102 | -.0093 | -0038 |
| | (-3.35) | (-1.64) | (-1.35) | (-.40) |

186

| | | | | |
|---|---|---|---|---|
| Ln crowding | .3006*** | .2718** | .0779 | .0367 |
| | (3.43) | (2.75) | (.67) | (.29) |
| % poor white or black families[a] | .0359* | .0420* | −.0017 | .0026 |
| | (1.99) | (2.22) | (−.14) | (.20) |
| Police per capita | ⋯ | −.0000 | | −.0001 |
| | | (−.06) | | (−.14) |
| Residential segregation | ⋯ | | .0214*** | .0194** |
| | | | (3.40) | (2.59) |
| South | .0542 | .0862 | −.1124 | −.0743 |
| | (.61) | (.93) | (−.87) | (−.43) |
| Northeast | ⋯ | −.2182 | ⋯ | −.1507 |
| | | (−1.49) | | (−.75) |
| North central | ⋯ | .0738 | ⋯ | .0705 |
| | | (.69) | | (.38) |
| Log likelihood | −80.96 | −78.25 | −112.40 | −111.27 |
| Chi-square test | 77.10*** | 82.53*** | 168.76*** | 170.85*** |
| Pseudo $R^2$ | .323 | .345 | .429 | .434 |

NOTE.—$t$-values are given in parentheses.

[a] When the dependent variable is white killings of whites, we calculate these indicators with whites; otherwise we use statistics on blacks.

* $P \leq .05$
** $P \leq .01$
*** $P \leq .001$

of whites or black killings of blacks, again because between-race antagonisms are unlikely to motivate within-race homicides. The findings support these expectations as well because they show that relationships between interracial competition and within-race killing rates are negligible. More generally, the contrasts in the results when we analyze the four possible combinations of within and between-race killing rates are consistent with theoretically based inferences that political and economic competition between the races influences white killings of blacks and black killings of whites. We find that racial conflicts over scarce resources spill over to alter interracial but not intraracial homicides.

REFERENCES

Balkwell, James W. 1990. "Ethnic Inequality and the Rate of Homicide." *Social Forces* 60.53–70.
Barth, F. 1969. *Ethnic Groups and Boundaries.* Boston: Little Brown.
Beck, E. M., and Stuart E Tolnay. 1990. "The Killing Fields of the Deep South. The Market for Cotton and the Lynching of Blacks, 1882–1930." *American Sociological Review* 55:526–39.
Beckett, Katherine. 1997 *Making Crime Pay: Law and Order in Contemporary American Politics.* New York: Oxford University Press.
Blalock, Hubert. 1967 *Towards a Theory of Minority Group Relations.* New York: Capricorn Books
Blau, Juddith R., and Peter M. Blau. 1982. "Metropolitan Structure and Violent Crime." *American Sociological Review* 47:114–28
Blau, Peter M. 1977 *Inequality and Heterogeniety.* New York: Free Press.
Blumer, Herbert. 1958. "Race Prejudice as a Sense of Group Position " *Pacific Sociological Review* 1 3–7
Bobo, Lawrence, and Franklin D Gilliam. 1990. "Race, Sociopolitical Participation, and Black Empowerment " *American Political Science Review* 84:377–93.
Bobo, Lawrence, and Vincent Hutchings. 1996. "Perceptions of Racial Group Competition: Extending Blumer's Theory of Group Position in a Multiracial Social Context." *American Sociological Review* 61:951–72.
Bonacich, Edna. 1972. "A Theory of Ethnic Antagonism: The Split Labor Market." *American Sociological Review* 37.547–59.
———. 1975. "Abolition, the Extension of Slavery, and the Position of Free Blacks: A Study of Split Labor Markets in the United States." *American Journal of Sociology* 37:547–59
Chiricos, Theodore G., and Miriam Delone. 1992. "Labor Surplus and Punishment: A Review and Assessment of Theory and Evidence." *Social Problems* 39:421–46.
Clark, Terry N. 1994. "Race and Class versus the New Political Culture." Pp. 21–78 in *Urban Innovation,* edited by Terry N. Clark. Thousand Oaks, Calif.: Sage.
Cleavor, Eldridge. 1968. *Soul on Ice.* New York: Dell-Delta, Ramparts.
Cohen, Lawrence C., and Marcus Felson. 1979. "Social Change and Crime Rate Trends." *American Sociological Review* 44:588–607.
Dickens, William, and Kevan Lang. 1985. "A Test of Dual Labor Market Theory." *American Economic Review* 75:792–805.
———. 1987 "Where Have All the Good Jobs Gone? Deindustrialization and Labor Market Segmentation." Pp. 90–102 in *Unemployment and the Structure of Labor Markets,* edited by K. Lang and J. Leonard London: Basil Blackwell.

Donohue J., and J. Heckman. 1991. "Continuous versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks." *Journal of Economic Literature* 29:1603–43

Farley, Reynolds. 1984 *Blacks and Whites—Narrowing the Gap?* Cambridge, Mass.: Harvard University Press.

Fishman, Mark. 1980. *Manufacturing the News.* Austin: University of Texas Press

Fosset, Mark A., and K. Jill Kiecolt. 1989. "The Relative Size of Minority Populations and White Racial Attitudes." *Social Science Quarterly* 70.820–35.

Gove, Walter R., Peggy Sullivan, and James Wilson. 1998. "How Are Income Inequality, Poverty, and Race Related to the Crime Rates and the Control of Crime?" Paper read at the Society for the Study of Social Problems meetings. San Francisco, August.

Greene, William H. 1993. *Econometric Analysis.* New York: Macmillan.

Hawley, A. H. 1950. *Human Ecology.* New York. Ronald Press.

Jackson, Pamela I. 1984. "Opportunity and Crime: A Function of City Size." *Sociology and Social Research* 69:172–93.

———. 1989. *Minority Group Threat, Crime, and Policing: Social Context and Social Control.* New York: Praeger.

Jacobs, David. 1979. "Inequality and Police Strength." *American Sociological Review* 44:913–25.

Jacobs, David, and Ronald E. Helms. 1996. "Toward a Political Model of Incarceration: A Time-Series Examination of Multiple Explanations for Prison Admission Rates." *American Journal of Sociology* 102·323–57.

——— 1997. "Testing Coercive Explanations for Order: The Determinants of Law Enforcement Strength over Time." *Social Forces* 75:1361–92.

Jacobs, David, and Robert M. O'Brien. 1998 "The Determinants of Deadly Force· A Structural Analysis of Police Violence." *American Journal of Sociology* 103:837–62.

Johnston, J. 1984. *Econometric Methods.* New York: McGraw-Hill.

LaFree, Gary D. 1982. "Male Power and Female Victimization: Toward a Theory of Interracial Rape." *American Journal of Sociology* 88:311–28.

Lane, Roger. 1997. *Murder in America.* Columbus: Ohio State Press.

Liska, Allen E., J. J. Lawrence, and M. Benson. 1981. "Perspectives on the Legal Order: The Capacity for Social Control." *American Journal of Sociology* 87:412–26.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, Calif.: Sage.

Maddala, G. S. 1977. *Econometrics.* New York: McGraw Hill.

Messner, Steven F., and Reid M. Golden. 1992. "Racial Inequality and Racially Disaggregated Homicide Rates." *Criminology* 30:421–45.

Messner, Steven F., and Scott J. South. 1992. "Interracial Homicide: A Macrostructural Opportunity Perspective." *Sociological Forum* 7.517–36

Mosher, Clayton, and John Hagan. 1994. "Constituting Crime in Upper Canada: The Sentencing of Narcotics Offenders, circa 1908–1953." *Social Forces* 72.613–41.

O'Brien, Robert M. 1987. "The Interracial Nature of Violent Crimes A Reexamination." *American Journal of Sociology* 92 817–35.

Olzak, Susan. 1990. "The Political Context of Competion: Lynching and Urban Racial Violence, 1882–1914." *Social Forces* 69 395–421.

———. 1992. *The Dynamics of Ethnic Competition and Conflict.* Stanford, Calif.. Stanford University Press.

Olzak, Susan, Suzanne Shanahan, and Elizabeth H. McEneaney. 1996. "Poverty, Segregation, and Race Riots: 1960 to 1993." *American Sociological Review* 58:590–613

Orcutt, James, and J. Blake Turner. 1993. "Shocking Numbers and Graphic Accounts: Quantified Images of Drug Problems in the Media." *Social Problems* 40:190–206.

Parker, Robert Nash. 1989. "Poverty, Subculture of Violence, and Homicide." *Social Forces* 67:963–1007.

Parker, Robert Nash, and Allan V. Horwitz. 1986. "Unemployment, Crime, and Imprisonment: A Panel Approach." *Criminology* 24:751–73.

Parker, Robert Nash, and M. Dwayne Smith. 1979. "Deterrence, Poverty, and Type of Homicide." *American Journal of Sociology* 85:614–24.

Price, James E. 1966. "Testing the Accuracy of the Crime Statistics." *Social Problems* 14:214–22.

Quillian, Lincoln. 1995. "Prejudice as a Response to Perceived Group Threat: Population, Competition, and Anti-Immigrant and Racial Prejudice in Europe." *American Sociological Review* 60:586–611.

———. 1996. "Group Threat and Regional Change in Attitudes toward African-Americans." *American Journal of Sociology* 102:816–60.

Roncek, Dennis W. 1992. "Learning More from Tobit Coefficients." *American Sociological Review* 57:503–7

Roncek, Dennis W , and Pamula A. Maier. 1990. "Bars, Blocks, and Crimes Revisited: Linking the Theory of Routine Activities to the Empiricism of 'Hot Spots.' " *Criminology* 29:725–53.

Rubenstein, Jonathan. 1973. *City Police.* New York: Farrar, Straus, and Girioux.

Sampson, Robert J. 1984. "Group Size, Heterogeneity, and Intergroup Conflict: A Test of Blau's *Inequality and Heterogeneity.*" *Social Forces* 62:618–39.

———. 1986. "The Effects of Inequality, Heterogeneity, and Urbanization on Intergroup Victimization." *Social Science Quarterly* 67:618–39.

———. 1987. "Urban Black Violence: The Effect of Male Joblessness and Family Disruption." *American Journal of Sociology* 93:348–82.

Sampson, Robert J., and Jacqueline Cohen. 1988. "Deterrent Effects of the Police on Crime." *Law and Society Review* 22:163–89.

Sherman Lawrence W., and Robert H. Langworthy. 1979. "Measuring Homicide by Police Officers " *Journal of Criminal Law and Criminology* 70:546–60.

South, Scott J., and Richard B. Felson. 1990. "The Racial Patterning of Rape." *Social Forces* 69:71–93.

South, Scott J., and Steven F. Messner. 1986. "Structural Determinants of Intergroup Association: Interracial Marriage and Interracial Crime." *American Journal of Sociology* 91:1409–30.

Stinchcombe, Arthur. 1963. "Institutions of Privacy in the Determination of Police Practice." *American Journal of Sociology* 69:150–60.

Taylor, Ian, P. Walton, and J. Young 1973. *The New Criminology.* London: Routledge & Kegan Paul.

Turk, Austen T. 1969. *Criminality and the Legal Order.* Chicago: Rand McNally.

Walton, John, and Charles Ragin. 1990. "Global and National Sources of Political Protest: Third World Responses to Debt Crises." *American Sociological Review* 55: 876–90.

White, Halbert. 1980. "A Heteroscedasticity-Consistent Covariance Matrix and a Direct Test for Heteroscedasticity." *Econometrica* 48:817–38.

Wilbanks, William. 1985. "Is Violent Crime Interracial?" *Crime and Delinquency* 31: 117–38.

Williamson, Joel. 1984. *The Crucible of Race: Black-White Relations in the American South since Emancipation.* New York: Oxford University Press.

Wilson, William J 1978. *The Declining Significance of Race.* Chicago: University of Chicago Press.

# Clocking Out: Temporal Patterning of Retirement[1]

Shin-Kap Han and Phyllis Moen
*Cornell University*

This article draws on life history data of the cohorts of recent U.S. retirees to examine the temporal patterning of retirement. Three major dimensions—historical context, social heterogeneity, and, most important, biographical pacing, measured by cohort, gender, and career pathway, respectively—operate simultaneously, yet unevenly, to affect various aspects of the retirement process. Findings suggest that changes over the past few decades have undermined the regularity in retirement timing that was a product of the convergence of diverse institutional features, anchored by a large core of men on traditional career tracks. Focusing on retirement, our model underscores the multiplex nature of the temporal structuring of the life course.

Two major trends characterize changes in retirement age over the past few decades, changes that have drastically altered its temporal frame. One is in the central tendency of its timing, and the other is in its variability. The downward shift in average age at retirement, at least for men in the West, has been amply documented (for reviews, see George [1993]; Guillemard and Rein 1993; Atchley 1982). This trend has been sustained and substantial, especially pronounced since the 1970s (cf. Quinn and Burkhauser 1994). Furthermore, as Kohli and Rein show (1991), it can be distinguished from a trend that gradually lowered the age of "normal" withdrawal from the labor force as pension systems were created and expanded from the late 19th century to the 1970s. It is described as "one of the most profound structural changes in the past 25 years" (Kohli and Rein 1991, p. 1).

But, perhaps of more import, the variability in the timing of retirement

has increased significantly during the same period. By the mid-1960s, age 65 had become the age around which retirement clustered in the United States, making it an effective reference point for a variety of purposes. Now, however, the transition is less clear-cut. The age-graded norm in retirement has become blurred, and the actual range of retirement age has expanded, making the transition "longer and fuzzier" (Kohli and Rein 1991; for reviews, see Settersten and Mayer 1997; George 1993; cf. Modell 1989; Rindfuss, Swicegood, and Rosenfeld 1987; Hogan and Astone 1986). It is becoming "destandardized" (Guillemard and Rein 1993), "deinstituti-onalized" (Guillemard and van Gunsteren 1991), and "age-irrelevant" (Neugarten 1979). As a result, what has long been regarded as the conventional age of retirement, age 65, is no longer "normal," or normatively prescribed (cf. Marini 1984).

The pairing between the lowering of average age at retirement and its increasing variability is peculiar, especially in the context of the historical trend of increasing institutionalization of life course transitions. Other than the fact that they have occurred concurrently, the relationship between the two is not self-evident and as yet is largely unexplored.

We consider the retirement transition to be a matter of contingent timing. Three major dimensions—historical context, social heterogeneity, and biographical pacing—are put forth, on which, we argue, retirement is timed. They act, either alone or in combination, upon a set of behaviors that affect when a person retires. Examining life history data of the cohorts of recent retirees, we identify and articulate these dimensions that pattern retirement timing. Our focus, in particular, is on the concept of biographical pacing and its role in timing retirement.

In the next section, we review the extant literature and elaborate upon a life course framework to formulate a multiplex time model of retirement. We discuss the three major dimensions involved in retirement timing and consider the ways in which these disparate, yet interacting, threads may be interwoven. Following a description of the data and measures, we identify a set of career pathway types. We then analyze the data and discuss findings within the framework of our proposed model of multiplex time. In conclusion, we locate the results in broader contexts and discuss several theoretical implications.

INSTITUTIONALIZATION AND DEINSTITUTIONALIZATION
OF RETIREMENT TRANSITION: A LIFE COURSE PERSPECTIVE

Until recently, historical developments in retirement had been in the direction of increased universality and decreased variability (Ransom and Sutch 1986; Anderson 1985; Graebner 1980). By the 1960s, retirement had become fully generalized in the United States as a normal feature of the

life course. Throughout industrialized societies, it became more prevalent and more tightly keyed to chronological age, establishing a standardized life course regime (Kohli and Rein 1991; Atchley 1982). Although the process affected primarily men's lives, women too have increasingly been incorporated into this regime with the rise in their labor force participation (Moen 1985). Societal regulation of retirement as such was primarily based on, and aligned with, a set of institutions—welfare systems broadly defined, including the state. Previous studies have thus identified a number of legal and regulatory changes (e.g., the legislation of the Social Security Act and Employees Retirement Income Security Act) and shifts in pension systems (e.g., the spread of private pension plans and the changing balance between public and private pensions) as determinants and conditions of the institutionalization of the retirement transition. They have depicted the modal transition pattern as reflecting conformity to normative timetables and constraints in opportunity and incentive structures (Guillemard and Rein 1993; George 1993; Kohli and Rein 1991; Mayer and Schoepflin 1989; Tuma and Sandefur 1988; Atchley 1982; Neugarten and Hagestad 1976).

Changes in recent years, however, have been so drastic in magnitude and variant in direction that the existing paradigm requires a critical reappraisal. The high degree of temporal regularity observed in the retirement regime of the 1960s and 1970s might have been more the exception rather than the rule. It was, we argue, a result of a multitude of factors coalescing and coinciding around a specific age, which in turn became defined as normative post hoc. The very definition of life course, "age-graded life patterns embedded in social institutions and subject to historical change" (Elder 1992, p. 1121), presupposes just such a contingent development of age-graded transition norms (see also O'Rand 1996; Mayer and Tuma 1990). Thus, undoing of the old regime, in which this extraordinary convergence around a single reference point occurred, seems to be a more natural progression than a startling deviation. Recent developments in the timing of retirement render visible the process in which these multiple threads, which were for a time brought together and kept in sync, are being pulled apart from one another.

A variety of factors have undermined the old regime. Three that pertain to the changes in the larger context—institutional, economic, and demographic—have primarily been the focus of previous research. First, there have been considerable changes in institutional arrangements around retirement. Oftentimes, however, they produce conflicting pressures. For example, whereas the age of eligibility for Social Security benefits has been lowered to 62 from 65, legislative measures have upwardly revised or eliminated mandatory retirement age. Another case in point: private pension funds encourage early exit, while recent public policy seems to

favor retarding the process (Guillemard and Rein 1993; Henretta 1992; Parnes and Nestel 1981). The confluence of policies and practices that had kept retirement timing in line seems clearly to be unraveling.

Second, the labor market itself is undergoing a period of fundamental transformation. During the past 15 years, in particular, many companies have sought to move away from the traditional model of mutual obligation, the "implicit contract," between employees and employers (Kalleberg, Knoke, and Marsden 1995; Osterman 1988). In the 1980s, for example, about 40% of the U.S. firms with more than 1,000 employees reduced their workforce through special early retirement incentives (Henretta and Lee 1996; Hardy, Hazelrigg, and Quadagno 1996; Guillemard and Rein 1993). Changes in the industrial structure at the aggregate level (such as the shift from manufacturing to service industries) have also affected the workforce (Jacobs, Kohli, and Rein 1991; DeViney and O'Rand 1988; Pampel and Weise 1983; cf. Sørensen and Tuma 1981). These economic transformations have made both employment more uncertain and retirement timing less predictable.

Third, we are witnessing a remarkable shift in the demographic base of the pool of actual and potential workers. Substantial improvement in life expectancy has lengthened the life span, suggesting the potential for extending employment to later ages (Settersten and Mayer 1997; George 1993; Tuma and Sandefur 1988). Age stereotypes regarding health, work performance, and productivity have also been significantly revised, and many people do indeed continue to work into later years or seek employment in some capacity following retirement from their "career" jobs (Kohli and Rein 1991; Maddox 1987). These developments have contributed to an expansion in the plausible range of retirement age.

All three of these changes have no doubt been crucial in bringing about the unfastening of what had come to be a "normative" retirement regime. The fact that these macrolevel transformations have been neither neatly coordinated nor closely synchronized adds yet another layer of variability in the deinstitutionalization of the retirement transition (Riley and Riley 1994; Kohli and Rein 1991).

These arguments are essentially based on structural and situational imperatives, where contextual changes lead to behavioral changes. Consider, however, the following evidence. First, it has been shown that financial incentives for early exit do not affect all social groups alike (e.g., Hardy et al. 1996; Guillemard and Rein 1993; Campbell and O'Rand 1988). Second, a series of analyses by Quinn and Burkhauser (1994) demonstrate that the impact of mandated retirement ages was much less than had been supposed (Quinn, Burkhauser, and Myers 1990, pp. 77–87). What these findings suggest is that contextual changes are necessary but not sufficient to account for life course institutionalization and deinstitutionalization.

Public policies, economic circumstances, and demographic profiles establish parameters, setting certain ground rules and stipulating available options. Yet how they actually translate into particular outcomes is an issue for further theoretical articulation and empirical research.

We propose two additional dimensions—"social heterogeneity" and "biographical pacing"—and incorporate them into our model of retirement timing. First, social heterogeneity refers to the fact that the impacts of societal changes are not uniformly distributed across the whole population. Foremost among the bases of such heterogeneity are the differences between men and women, which have been extensively documented (Moen 1996a, 1996b; Henretta, O'Rand, and Chan 1993; Kohli and Rein 1991; DeViney and O'Rand 1988; Pampel and Park 1986; Henretta and O'Rand 1983; Streib and Schneider 1971).[2] These differences reflect the gendered opportunity structure as well as variations by gender in expectations and preferences. The process of choosing "the right time to retire" has much to do with reference group dynamics. Individuals' frames of comparison are not typically based on the experiences of a random sampling of other people or an averaging of overall population (Hardy et al. 1996). Rather, comparisons tend to be contextual. Gender provides one of the most salient—easy and clear—comparative frames, inducing homogeneity within gender and fostering distinctive norms between men and women. Yet the structural and cultural contexts of comparison are themselves in flux. For example, while women's work history has historically been characterized by tangential and transient ties to employment, their labor force participation rate has continually increased, and the gap between men's and women's workforce experience has been steadily closing (Tomaskovic-Devey 1993; Guillemard and Rein 1993; Moen 1985; Waite 1981).

The second dimension we incorporate is biographical pacing. Explanations based exclusively on exogenous factors ignore the fact that life course transitions are a product of individual action as well as institutional and historical forces. As social regulation weakened, the tension between "biography" and "history" (Campbell and O'Rand 1988; also see George 1993; Mayer and Tuma 1990; Mayer and Schoepflin 1989; Barley 1989; Dannefer 1988; Elder 1985) has increasingly come to the fore. In other words, the loosening of institutionalized retirement regime was coupled with increased voluntarism and individuation, rendering the endogenous, subjective aspect of life course processes more salient (cf. Modell 1989; Mayer 1986). More concretely, on the one hand, the range of possible choices in retirement timing decisions has considerably broadened due to

[2] Age and cohort differences have also been noted in previous research (e g , Henretta and Lee 1996; Hogan and Astone 1986).

the macrolevel transformations discussed above. On the other hand, in tandem with and because of this changing landscape of constraint and opportunity, more and more workers have begun to leave under terms of their own choice and at their own pace. To date, this aspect of choice within constraint is the least articulated in the retirement research literature.

## TIMING RETIREMENT ON MULTIPLE CLOCKS: THE MODEL

To sum up, the retirement regime is being reorganized, if not undone.[3] This reorganization presents "perhaps the most explicit challenge" for life course theory, not only as an explanatory task but also as a theoretical window on the construction of the life course in general and the articulation of endogenous and exogenous factors underlying the life course in particular (Guillemard and Rein 1993, p. 1). Dannefer's (1988, p. 374) call for theoretical language describing the phenomenon as a systematically stratified and differentiated process is, thus, an apt and timely one. The idea that life course transitions and trajectories are subject to more than one temporal dimension is not new. Previous research has suggested that tempo and scheduling of life course in general, and retirement timing in particular, are shaped by multiple factors and alignment between them (Mayer and Schoepflin 1989; Featherman 1986; Campbell and O'Rand 1988). This research called for a "multilevel and multitime framework" (Mayer and Tuma 1990, pp. 6–7; Elder 1985; see, e.g., Petersen and Spilerman 1990). What we question is the validity of a presumption—frequently, though implicitly, made in the past—that various temporal dimensions are operating in tandem, producing a high degree of social regulation in life course transitions (Mayer 1986).

Recent changes in the transition to retirement provide us with a strategic site to probe into the multiplex nature of temporality associated with life course institutionalization and deinstitutionalization. The retirement decision is for most workers primarily a timing problem, that is, they are trying to pick "the best time to retire" (Hardy et al. 1996). That decision, however, is both biographically embedded and historically conditioned. It is also, as noted earlier, becoming quite heterogeneous (Campbell and

[3] The deinstitutionalization of the retirement regime is evident not only in timing but also in the form of retirement. Alternatives to the traditional, i.e., clearly demarcated, determinate, orderly, absorbing, definitive, and "crisp," form of retirement have multiplied. As a result, it has become more variable, diffuse, imprecise, and contingent (Mutchler et al. 1997; O'Rand 1996; Elder and O'Rand 1995; Guillemard and Rein 1993; Henretta 1992). The notion of retirement as an all-or-nothing dichotomy is no longer adequate, and reconceptualizing it as a process instead of an event has already been suggested in previous research (Atchley 1982).

O'Rand 1988). These constitute the three major temporal dimensions we observe in retirement: (historical) context, (social) heterogeneity, and (biographical) pacing.

First, we have described the way macrolevel environment—mainly the state and the labor market—and historical circumstances set aggregate baseline parameters around retirement transition. These are external influences, consisting of long-term trends and random fluctuations. We call this "historical context." Second, individuals as purposive actors make the decision to retire. On the one hand, courses of action around retirement can be volitional and internal. Yet, on the other hand, they are bound by cumulative contingencies. That is, there is a narrative logic in one's career that can go on its own momentum, setting its own pace. The product of this tension is what we call "biographical pacing." Finally, between these two are social groupings, such as gender, where we typically locate normative expectations (Henretta 1992; O'Rand 1996; Elder and O'Rand 1995). We call this "social heterogeneity." Note, however, that this is not an independent temporal dimension per se. Rather it constitutes a major source of variation, since the ways in which historical context and biographical pacing operate differ across social boundaries such as gender.

Pacing Life Course Transition: Career Pathways and Retirement

Among the three temporal dimensions in our model, biographical pacing has been the most neglected in previous research. The concept is based on a central tenet of the life course perspective: to understand behavior at any one life stage requires knowledge of prior transitions and trajectories (Elder 1992, 1995; O'Rand and Henretta 1982; cf. Hughes [1937] 1994). Transitions are always embedded in the trajectories that give them distinctive forms and meanings (Elder 1995). Retirement, therefore, should be viewed as a transition occurring within the context of overall career trajectory, reflecting biographical pacing, which affects the timing of this key transition.

Career pathways are typically thought of as a shorthand for occupational mobility. Yet they can be conceptualized and operationalized in a variety of ways (Breiger 1995; Rosenfeld 1992).[4] First, career pathways may differ in terms of level of overall mobility. Taking career pathways to be series of positions, one may expect an orderly and hierarchical progression of jobs (Wilensky 1961) or individual achievement (Spenner, Otto, and Call 1982) to be the norm. The extent and shape of deviation

---

[4] We employ the term "career" neutrally with respect to orderliness and disorderliness of work history as in Spilerman (1977).

from this norm are of crucial importance, however, and need to be examined empirically (e.g., Rosenbaum 1984; Spilerman 1977). Variations in career pathways can be expected to affect other behaviors, such as retirement. For instance, one would anticipate that those having experienced uneven, or downward, career pathways to be less likely to plan for retirement and more likely to have greater variation in their retirement timing.

Second, career pathways may similarly vary in terms of their relative continuity in employment. For instance, workers who have frequently moved in and out of the labor force, in and out of employment, or in and out of jobs in many firms, may be less invested in their careers and less attached to the labor force (Wilensky 1961), viewing early retirement incentives and possibilities as a relief, a means of formally exiting what have been erratic, disrupted careers (Moen 1996b). Also germane is the importance of timing. For instance, workers in their sixties who had planned on retiring anyway should be less likely to view being laid off due to downsizing as unfinished business, and consequently less likely to seek subsequent employment. By contrast, workers in their early fifties may view the same event quite differently.

In sum, various characteristics of career pathways, we argue, bear upon the behaviors before, around, and after the retirement. We focus, empirically, on the labor force experiences of retired men and women, identifying and describing their career pathways to retirement. Taking into consideration overall pattern of trajectory as well as important transitions within it, we investigate whether and how these experiences influence various aspects of retirement.

Our model depicts retirement as a set of behaviors temporally patterned on the three proposed dimensions. It is hypothesized to be timed on three separate clocks, all ticking simultaneously yet to different beats (Mayer and Tuma 1990; Dannefer 1988). Note that we specify retirement as a set of behavioral variables, permitting us to examine various aspects and stages that constitute retirement, from retirement planning to postretirement employment. In other words, retirement is seen as a process, rather than a one-time, one-way exit from the labor force. (See n. 3, above, for further discussion.)

The model generates a number of novel research questions. Do these three factors—historical context, gender difference, and career pathway—exert separate, additive influences on planning and timing of retirement and postretirement employment? Or do they interact with one another? Do these effects operate in the same direction? Is any one temporal dimension primary in timing retirement? Do these clocks operate consistently across various behavioral aspects of retirement? Do any of them operate at cross-purposes? Focusing on intertemporal patterning of life course, we address these questions in the following analysis.

## DATA AND MEASURES

### Data

We analyze data collected in the first wave of the Cornell Retirement and Well-Being Study (CRWB). The respondents are 304 older workers and 458 retirees from six large manufacturing and service companies of four cities in upstate New York who were ages 50–72 at the time they were interviewed in 1994–95. Note that these six companies, from which we collected our data, do not represent the broad spectrum of all companies in the United States. Rather, they belong to the upper tier of the spectrum. Two of them are among Fortune 500 companies and the other four are the largest in the state in their respective industries. One might, hence, consider that our respondents are largely representative of those who have experienced the best of circumstances in their career and in their retirement. Yet, these are also the companies that spearheaded drastic restructuring and downsizing in the 1980s and 1990s. Respondents were randomly selected from the lists provided by their employers and initially contacted by letter and telephone to request their participation and arrange for an interview.[5] The interviews, ranging from one to two and a half hours, were conducted face to face, save for those who had relocated to new communities, which were conducted by telephone.

The principal survey instruments include a structured interview schedule and a booklet of self-administered questions drawn from a number of sources, including the Health and Retirement Survey (Juster 1992) and the Quality of Employment Survey (Quinn and Staines 1979). These instruments were extensively pretested prior to administration. Of special interest to us is the respondents' employment history, which we draw from the collection of detailed life history data.

In the analysis that follows, we use data only on the retiree subsample. There are 212 women (46%) and 246 men (54%), with an average age of 63 years, who have spent anywhere from one month to more than 19 years in retirement. Most retired around 1990. Because retirement is increasingly distinct from a permanent exit from the workforce, we operationalize being "retired" as receiving a pension (or retirement package) from one of the six companies.[6] Respondents had been last employed in a wide

---

[5] The overall response rate was 78% among those contacted.

[6] Note that this is not the only way to define retirement status. Parnes and Less (1985; Parnes and Nestel 1981) have examined three alternative operational criteria of retirement—(1) subjective: individual's own perception of having stopped working at a regular job; (2) income: receipt of Social Security or other pension income; and (3) participation: partial or complete withdrawal from the labor market (e.g., working less than 1,000 hours during the year)—and found that retirement status of four-fifths of the men could be defined unambiguously. Such a definition depends also on whether

range of preretirement jobs spanning much of occupational hierarchy. To account for the possible biases in selecting only the retirees from the initial sample ("incidental truncation"), and to ensure consistent estimates, we adopt Heckman's (1979; see also Greene 1993; Winship and Mare 1992) two-equation model and correct for selection bias in the model estimation.

## Dependent Variables: Multiple Measures of Retirement

In addition to the age respondents retired, we gathered data on a set of four variables that are pertinent to retirement timing. Two of them have to do with various aspects of the retirement age norm: age they began planning retirement and age they expected to retire, or their target retirement age. We also have information on whether the respondent had taken an early retirement incentive (ERI), which directly affects retirement timing. Finally, we examine respondents' postretirement employment, which could be linked to their retirement planning, expectations, and timing (cf. Ekerdt and DeViney 1993). Each of these multiple measures bears on various aspects and stages of retirement, allowing us to examine the composite and processual nature of retirement.

## Explanatory Variables

Based on historical events and the ways in which these respondents collectively experienced them (Elder and Pavalko 1993; Burt 1991; Hogan 1978; Ryder 1965), we delineate three cohorts to capture historical context: (1) born before 1929, (2) born between 1929 and 1934, and (3) born after 1934. By the time of interview (1994–95), they were (1) over 66 years of age, (2) 60–65, and (3) less than 60 years of age, respectively. Benchmark events in their life course are the Great Depression at one end and massive corporate restructuring in the late 1980s and early 1990s at the other. The latter, in particular, bears on the changing opportunity structure, which could have affected these three cohorts differentially given that they were at different life, and career, stages in the late 1980s and early 1990s (Kalleberg et al. 1995; Kotter 1995; Kanter 1989; Osterman 1988). By 1985, the onset of massive corporate downsizing, they were at ages 56 and over, 50–55, and under 50, respectively (see table 6, below).

The main locus of social heterogeneity we consider is gender. We have oversampled women to balance the data in terms of gender, making possi-

---

"retirement" is conceived as an event (i.e., "once retired, always retired") or a current status (i.e., "What is the individual doing in a particular year?"; see also Atchley [1982]).

ble statistically reliable estimates of gender differences. Note that, since our respondents are from the six companies mentioned above, the data represent only the experiences of men and women who had been employed and retired. What our findings show, hence, might be closer to the national average picture for men but might be rather skewed for women in the general population (in that only those women who had been employed and retired are included in our sample, while those who had been exclusively full-time homemakers, for instance, are not).

To address whether or not, as well as how, retirement timing is influenced by prior work experiences—what we call biographical pacing—first requires definition and measurement of those experiences. Through an examination of the life history data detailing various aspects of career transitions and trajectories over the life course leading to retirement, we chart a set of typical career pathways for retired men and women. We discuss the theoretical issues and operational details involved in the next section.

Two main control variables pertain to the respondents' human capital, education (years of schooling completed), and tenure (years in the organization retired from) (Becker [1964] 1993; Mitchell and Fields 1982; Mincer 1974). The latter also serves as a rough proxy for eligibility and amount of pension, thus, at least indirectly, taking into account the financial aspects of retirement timing (Hardy et al. 1996; Guillemard and Rein 1993, p. 479). The selection bias estimate ($\lambda$) is obtained from a probit regression model (Greene 1993), as shown in table A1 in the appendix, based on gender, cohort, and two main control variables: education and tenure. Dummy variables for the companies the respondents are selected from are also included in the model. The selection bias is controlled for in all estimations reported below. A few additional variables related to the circumstances around retirement are included in estimations of the likelihood of postretirement employment, which will be discussed later.

## CHARTING CAREER PATHWAYS

To chart the regular patterns in employment history, we employ a sequence analysis technique known as "optimal matching" or "optimal alignment." It is a new method for old ideas; sequences of events or phenomena have been a concern of a wide variety of research in social sciences (Abbott 1995a). Life course researchers, in particular, have been interested in this issue over the past two decades, for it is at the core of life course perspective's two key constructs—trajectories and transitions—both conceptually and methodologically (Pavalko 1995). However, the methods developed and used thus far have mostly focused on individual events, not the sequence as a whole, that is, sequence qua sequence.

Event history analysis is a typical case in point, where individual spells are the unit of analysis, with transitions between spells the events of interest (Tuma and Hannan 1984). By contrast, we consider each respondent's whole sequence, examining the overall patterning of career trajectories. Specifically, we take into account the incidence, timing, and duration of diverse events, and their sequence across multiple domains of career (cf. Pavalko, Elder, and Clipp 1993; Rindfuss et al. 1987; Rosenbaum 1984; Hogan 1978; Wilensky 1961).

The data on employment histories of retirees provide information on transitions and trajectories over the life course in occupation, work status, and organization from age 30 until retirement. Occupation was coded with a nominal coding scheme used in the CRWB study ($k = 64$; see table A2 in the appendix). Work status was coded into 5 categories (employed full time, part time, alternating between the two, sporadic/seasonal, and not employed). And, last, organization was coded by counting the number of companies or employers the respondent has worked for up to age $t$. Among the retirees, reconstruction of complete employment histories was possible for 401 cases.[7] Using yearly interval as unit-time, the data were transformed into sequence data format, that is, strings of codes. Presented in figure 1 is an example of a typical female retiree, Katie. Having been out of the labor force for a long time (work status = 5, occupation = 64, organization = 0), Katie started working full-time at age 37, preparing and serving food (1, 42, 1). After three years, at age 40, she moved to another company to work, again full-time, as a machine operator (1, 60, 2), from which she retired after 19 years of continuous employment. This serves as an illustration of strings that contain information on incidence, timing, duration, and sequence for each of the three intersecting dimensions of employment history.

Based on the interelement distance matrices, specifying pairwise distance between codes (see tables A3 and A4 in the appendix), the optimal matching algorithm produces measures of dissimilarity between these sequences.[8] Three dissimilarity matrices—occupations, organizations, and

[7] With regard to the baseline variables (refer to table A1 in the appendix), there were no significant differences between those who were retained (87.6%) and those who were not (12.4%).

[8] See Abbott and Hrycak (1990) for an extended introduction to this technique, and Abbott and Barman (1997); Stovel, Savage, and Bearman (1996); Blair-Loy (1999); and Chan (1995) for its recent application in substantive areas. For optimal matching of sequences for this analysis, we used an adapted version of DISTANCE by Stovel (1996) written in SAS/IML. It is based on the same algorithm described in Abbott and Hrycak (1990). It allows, however, one to analyze data with larger $N$ (> 150; cf. Abbott 1995b).

| Age ($t$) | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Occupation: | 64 | 64 | 64 | 64 | 64 | 64 | 64 | 42 | 42 | 42 | 60 | 60 | 60 | 60 | 60 | 60 |
| Organization: | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 |
| Work Status: | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

| Age ($t$): | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | ... |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Occupation. | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | R | | ... |
| Organization: | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | R | | .. |
| Work Status: | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | R | | . . |

FIG. 1.—Sequence data: an example For occupational category code, 64 = not employed (unemployed/out of labor force); 42 = food preparation and service; 60 = machine operators, assemblers, and inspectors. For more details, see table A2 in the appendix. Organization indicates $N$ of companies worked for from age 30 to age $t$. Work status category code, 1 = full-time; 5 = not employed (unemployed/out of labor force). $R$ denotes retirement.

work status—are then normalized and combined to form one overall matrix.[9] Hierarchical clustering is performed on this combined dissimilarity matrix to identify a set of typical pathways that simultaneously takes into account all three dimensions. This process of identifying typical career pathways is analogous to that of using "structural equivalence" as a basis of grouping actors in network analysis (Burt 1983; White, Boorman, and

[9] In handling multiple dimensions, our strategy is slightly different from that employed by Abbott and Hrycak (1990), Blair-Loy (1999), and Stovel et al. (1996). Whereas they start with one string combining information on all dimensions, we combine the dimensions later Nonetheless, both obtain distances additively, thus are equivalent to each other. As the number of dimensions increases and as the number of "elements" increases for each dimension, however, the former strategy becomes quickly unwieldy. With our data, for instance, we could have defined 3,520 elements (5 for work status × 11 for organization × 64 for occupation), and our full substitution cost matrix might have been 3,520-by-3,520. We have, instead, opted for the second strategy, which is much more manageable.

TABLE 1

ANOVA on Pairwise Distances within- and between-Clusters

| | N | Number of Pairs* | Mean Distance | SD of Distance | P of Difference† |
|---|---|---|---|---|---|
| *Within*-cluster pairs:‡ | | | | | |
| Cluster 1 ... ..... ................ . | 46 | 1,035 | .361 | .089 | < 001 |
| Cluster 2 . ............... .. .. ... | 154 | 11,781 | .409 | .095 | <.001 |
| Cluster 3 ... ............ .. . .. ... | 160 | 12,720 | .468 | .093 | <.001 |
| Cluster 4 ........... . .... ........... | 10 | 45 | .516 | .086 | 116 |
| Cluster 5 ........ ................ . | 21 | 210 | .375 | .122 | < 001 |
| All *within*-cluster pairs§ ... ... | 391[1] | 25,791 | .436 | .100 | <.001 |
| All *between*-cluster pairs ... ... | | 50,454 | .539 | .100 | |
| All pairs .................. .. .......... | | 76,245 | .504 | .111 | |

* Number of pairs is obtained by $N(N - 1)/2$
† Significance based on the least significant difference (LSD) pairwise multiple comparison test
‡ With all between-cluster pairs, $F_{(5,76239)} = 4340.0$, $P < .001$, $\eta^2 = .222$.
§ With all between-cluster pairs, $F_{(1,76243)} = 18027.6$, $P < .001$, $\eta^2 = 191$.
[1] Excludes 10 cases not clustered into the five clusters With those 10 cases included, virtually identical results are obtained, $F_{(7,80192)} = 3385.2$, $P < 001$, $\eta^2 = .228$; $F_{(1,80198)} = 20081.7$, $P < .001$, $\eta^2 = 200$.

Breiger 1976). In other words, the career pathways clustered together into a type are "sequentially equivalent" to one another. We discerned in the data five distinct clusters, based on the criteria of *F*-ratios and other tests as shown in table 1.[10] We call these "career pathway types," or "pathway types" for short. They are labeled as "delayed entry pathway" (cluster 1), "orderly pathway" (cluster 2), "high-geared pathway" (cluster 3), "steady part-time pathway" (cluster 4), and "intermittent pathway" (cluster 5), respectively, for easy identification and reference. The labels describe major characteristics observed in each type (cf. Moen 1985; O'Rand and Henretta 1982).

Five Pathway Types

In the first column of figure 2, we illustrate how the respondents have experienced their careers in terms of occupational status and mobility by assigning a socioeconomic index (SEI) score to each occupational category (Nakao and Treas 1994). The vertical coordinate indicates occupational

[10] The solution we adopted here is based on an initial solution with seven clusters. Two of the clusters, however, are dropped from the analyses reported below due to the small *N*s (seven and three), leaving five clusters.
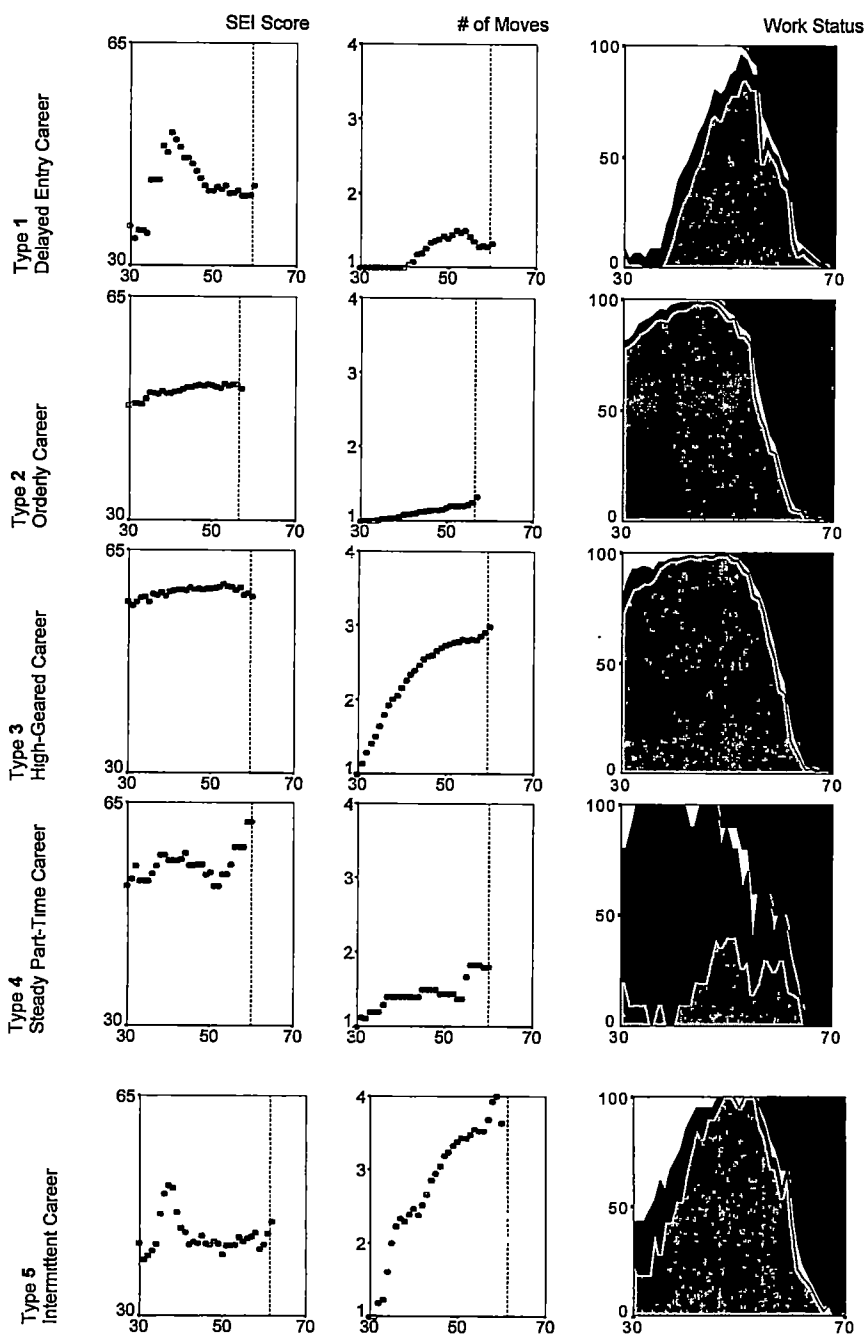
FIG. 2.—Age profiles for career pathway types. The dashed vertical lines indicate the median retirement age of respondents in each pathway type. For "SEI Score," average assigned score is used (see table A2). "# of Moves" indicates the average number of companies or employers worked for up to the age indicated on the horizontal axis. For "Work Status," the white area represents the proportion of respondents unemployed or out of the labor force, the light grey area, those employed full-time; the black area, those employed part-time; the dark gray area, those retired.

prestige in SEI score, and the horizontal coordinate indicates age. The dots represent the average SEI score across age.[11] The dashed vertical line indicates the age where half of the respondents in each pathway type have retired (i.e., median retirement age). In terms of average SEI scores, high-geared pathway type—consisting mostly of those in executive and managerial positions—is the highest, while delayed entry and intermittent types—primarily made up of those in clerical positions—are the lowest (see table 2 for more details). In terms of change and status mobility, the contrast is between orderly and high-geared pathway types and those of delayed entry, steady part-time, and intermittent types. While the former show a steady and continuous upward mobility, the latter reflect quite a bit of fluctuation.

The second column of figure 2 shows how the respondents have moved between companies or employers. The vertical coordinate indicates average number of companies or employers worked for up to that age. With respect to interorganizational mobility, the basic contrast is between high-geared and intermittent pathway types, "movers," and those of delayed entry, orderly, and steady part-time types, "stayers." There is a subtle, yet important, difference among the movers, however. For the high-geared pathway type, mobility continues until the midforties, then tapers off afterward. For the intermittent type, by contrast, mobility remains high until retirement.

Respondents' trajectories with respect to their work status over time are shown in the last column. (See table 2 for the average amount of time spent in each status over the entire period.) The white area represents proportion of respondents unemployed or out of the labor force, the light gray area those employed full time, the black area those employed part time, and the dark gray area those retired. There are three distinct groups based on work status profile over the life course: full-time group for orderly and high-geared pathway types, part-time group for steady part-time types, and another group consisting of delayed entry and intermittent types. The last group has an extended period of being out of the labor force early on, suggesting that these respondents start their work careers late. This group also shows a significant presence of part-timers.

Overlaying the three columns of graphs on top of one another permits more elaborate descriptions. The high levels of interorganizational mobility found in high-geared and intermittent pathway types turn out to be based on two entirely different dynamics. For the former, changing jobs across organizations seems to be a way of achieving upward mobility,

[11] The score is assigned by mapping the 1990 census three-digit codes onto the occupational categories we use as shown in table A2 in the appendix. Average SEI scores are calculated only for those who are employed at the time.

## TABLE 2

### FIVE PATHWAY TYPES AND THEIR CHARACTERISTICS

| | Delayed Entry Career (1) | Orderly Career (2) | High-Geared Career (3) | Steady Part-Time Career (4) | Intermittent Career (5) |
|---|---|---|---|---|---|
| N ......... | 46 | 154 | 160 | 10 | 21 |
| Gender composition: | | | | | |
| (% men) ......... | .0 | 64.9 | 61.9 | 30.0 | .0 |
| (% women) ......... | 100 0 | 35.1 | 38 1 | 70 0 | 100.0 |
| Education (mean/years) .. | 12.64 | 13 25 | 14.61 | 13.40 | 12.86 |
| SEI (mean) .. ......... | 42.8 | 50.3 | 58.5 | 54.6 | 42.5 |
| No of organizations (mean)......... | 1 5 | 1 2 | 2.8 | 1.7 | 3.7 |
| Work status. | | | | | |
| Full-time (%) ......... | 47.4 | 94.6 | 94.4 | 26.2 | 73 7 |
| Part-time (%) ......... | 9.3 | 1.5 | 2.3 | 69 6 | 10.6 |
| Unemployed/OLF (%) ......... | 43.4 | 3 9 | 3.3 | 4.2 | 15.7 |

NOTE.—For each work status, percentages are obtained by [(person-years in that work status)/(total person-years)] × 100. The differences among the five pathway types in terms of the criterion variables are tested For gender composition, likelihood-ratio ($L^1$) test is used, and for the others, $F$-test is used The tests are significant at $P < 001$

while for the latter, job shifts across organizations are merely an indication of instability, reflecting frequent exits and reentries. The comparison between orderly and high-geared pathway types provides another intriguing case. Both experience upward mobility, but that of the orderly type seems to be based on "ladder climbing" in the same firm, that is, moving up through the internal labor market. This is more or less what the "ideal" type of employment history, as constructed in the Unite States during the 1950s, 1960s, and 1970s, would look like, that is, continuous employment in a single organization for the entire span of a career (Wilensky 1961). By contrast, the mobility pattern captured in the high-geared pathway type seems to be based on "ladder hopping" (Kanter 1989; Kotter 1995).

Examining characteristics of respondents sorted into the five pathway types makes their distinctiveness more apparent. Gender is a crucial factor that we expect to shape many aspects of career pathways. Almost all the men are found either in orderly or high-geared pathway types. While these two are predominantly male, there is a significant presence of women as well (35.1% and 38.1%, respectively). Delayed entry and intermittent career pathway types consist exclusively of women, while the steady part-time type is predominantly composed of women (see table 2). In short, we observe two patterns in our data. First, on average, there seem to be distinct and separate career pathways for men and women ($L^2_{(4)} = 117.159$; $P = .000$). Yet, second, working women seem to have traveled quite diverse paths, whereas men's career paths tend to be much more standardized, falling primarily into a couple of career pathway types.[12]

One puzzling issue noted in reading figure 2 can be resolved with this information on gender composition. On the one hand, the peak in SEI score for delayed entry type occurs around age 40, when only a small proportion of the respondents in that pathway type are employed full time. On the other hand, a massive influx into the labor force takes place right after that point, which is accompanied by a rapid decline in SEI score. Intermittent type shows a similar pattern, with slightly different timing. The common thread between these two pathway types is their gender

[12] Also note that, on the one hand, we find that women's work patterns have increasingly taken on the appearance of men's—i.e., full-time permanent attachment—as seen in the sizable presence of women in orderly and high-geared career pathway types (Masnick and Bane 1980; Kreps and Clark 1975; Waite 1981; Moen 1985). On the other hand, there is also evidence that men—some, at least—have been experiencing changes in the social organization of work and family of the last several decades by taking on the patterns that used to be associated with women's careers, as shown in the steady part-time type. Both illustrate that the issue of gendered careers cannot be addressed simply by contrasting stylized men's vs. women's career paths. Rather, one needs to have a more refined perspective on the differentiation between as well as among men and women within particular historical contexts.

composition, that is, they both consist exclusively of women. Most women in both pathway types are not employed in their thirties because they are spending full time having and raising their children; but they return to paid work in their forties. However, they return to jobs with little prospect of upward mobility.

Level of schooling is, also as expected, related to the manner in which respondents are sorted into various career pathway types (see table 2). This is partly due to the confounding of gender and education effects: men are, on the whole, more educated than women in this sample, as in the general population of this age group. Yet, even after controlling for gender (not shown), the five types show substantial differences in average level of education. Those in the high-geared career pathway type are the most highly educated, followed by orderly and steady part-time types. Delayed entry and intermittent pathway types are traveled mostly by high school graduates, whereas high-geared type is a path common to college graduates. Among women, those following high-geared paths are the most educated; there is little difference in education among women in the other four types.

In sum, the orderly pathway type seems to represent what has been thought to be the "ideal" career path, that is, stable, continuous, and upwardly mobile. Those on the delayed entry pathway type, exclusively women, enter the workforce after their childbearing years. Although they work typically at low SEI jobs, these jobs are relatively stable. The high-geared type characterizes the experience of those who are highly educated and upwardly mobile. They start off high on the occupational ladder and move about quite a bit. The steady part-time career pathway type consists of a small group of people working mostly part time. Yet they show low levels of interorganizational mobility and are relatively successful in terms of SEI score and upward mobility. The intermittent pathway type consists exclusively of women and is the least stable of all. Although it shares many of the characteristics of the delayed entry type, it distinguishes itself by a trajectory of higher mobility across organizations, mostly due to frequent exits and reentries.

ANALYSIS AND FINDINGS

In this section, we test multivariate models of retirement timing. We investigate not only the effects of the three dimensions—historical context, social heterogeneity, and biographical pacing—but also the changing patterns of those effects across the five dependent variables characterizing the retirement transition. We address the following questions: Do respondents differ in how and when they retire according to the historical contexts in which they made the retirement decision? Does gender shape the ways in

which respondents end their working careers? Do respondents traveling varying career pathways pace their retirement transition differently?

## Retirement: Planning, Expecting, and Doing It

First, we find that respondents' age at retirement varies considerably, yet most retired in their late fifties and early sixties (median: 59; IQR: 56–62; see fig. 3). But are respondents retiring when they expected to? To address this issue we examine two variables. One deals with when they began planning their retirement, and the other specifically taps into target retirement age if they did plan. Note that these two are subject to retrospective biases. In addition to the recall problem, there is a possibility of consequence affirming rationalization, given that the respondents have already retired.[13] Nonetheless, responses to these two questions reflect underlying norms regarding retirement timing in this sample.

Although correlations between the three are quite substantial, especially between target and actual retirement ages, each varies distinctively, as shown in figure 3.[14] A few respondents report having begun planning their retirement very early, yet the pace begins to pick up in earnest around age 40, reaching a peak around 50. Some seem to have begun rather too late to really "plan ahead." Among the three variables, this is the one with the greatest dispersion, suggesting lack of strong normative consensus on when to begin planning for retirement. When they did plan, respondents report target retirement ages that are easily markable, such as 55 or 65. Target retirement age ranges from the late fifties to early sixties. Age 62 is the most frequent choice, obviously conditioned by the Social Security benefit eligibility requirement.

Table 3 shows, first, that men's timing is earlier than women's on all three measures, particularly in the age they began planning. This might be due to the fact that most women in our data entered the workforce later than men, that women are less likely than men to be able to afford retirement, or that women are less adequately socialized regarding the

[13] Research shows that the reliability of expectation as a predictor of actual retirement is related to the age at which the expectation is measured (Nestel 1985). For instance, men who are relatively young when asked their target retirement age report earlier ages, on the average, than older men.

[14] The correlation coefficients between the three are:

| Age | Began Planning | Target |
|---|---|---|
| Began planning | | |
| Target | .366 $(P = .000)$ | |
| Actual | .388 $(P = .000)$ | .683 $(P = .000)$ |
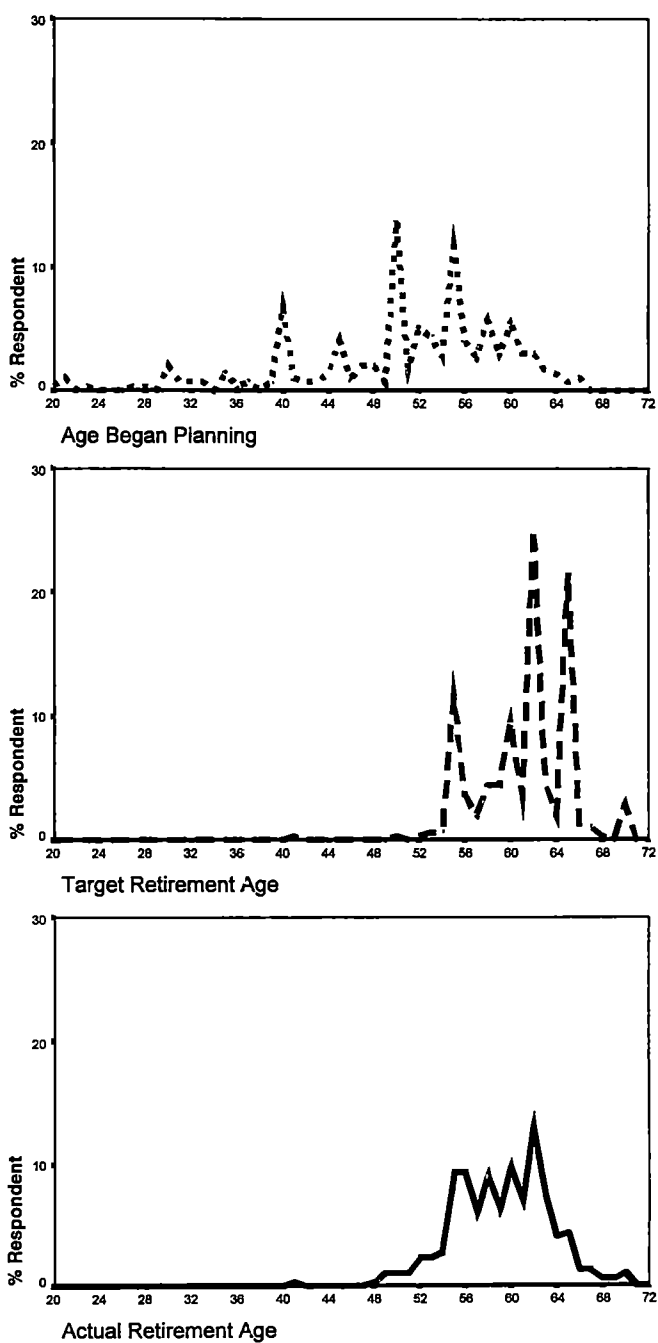
FIG. 3.—Timing retirement (CRWB study, wave 1 [1994–95]), $N = 458$

TABLE 3

Time Retirement by Cohort, Gender, and Pathway Type

| | Age Began Planning (1) | Age Expected to Retire (2) | Age Retired (3) | (3) − (2) |
|---|---|---|---|---|
| Gender: | | | | |
| Men | 48 54 | 60 84 | 58.72 | −2.21 |
| | (9.51) | (4.01) | (4 43) | (3 38) |
| Women | 53.96 | 61.65 | 59 57 | −2.19 |
| | (7.52) | (3.89) | (3 88) | (3.05) |
| F-test | P = .000 | P = 043 | P = .043 | P = .947 |
| Cohort: | | | | |
| 1 (1923–28) | 54.17 | 63.26 | 61.83 | −1.57 |
| | (8.92) | (3.40) | (3.55) | (2.86) |
| 2 (1929–34) | 50.22 | .60 70 | 58.96 | −1 77 |
| | (8.48) | (3.19) | (2.68) | (2.92) |
| 3 (1935–43) | 46.16 | 58.31 | 53 99 | −4.34 |
| | (8.28) | (4.32) | (2.84) | (3.63) |
| F-test | P = .000 | P = .000 | P = .000 | P = .000 |
| Pathway type | | | | |
| 1 (delayed entry) | 56.73 | 61.74 | 59.91 | −2.00 |
| | (4.81) | (3.51) | (3.92) | (2.49) |
| 2 (orderly) | 48.89 | 60.43 | 57.75 | −2.81 |
| | (9.31) | (4.06) | (4.15) | (3.69) |
| 3 (high-geared) | 50.80 | 61.77 | 60.17 | −1.67 |
| | (9 49) | (3.65) | (3 68) | (2.91) |
| 4 (steady part-time) | 57.60 | 60 50 | 59.00 | −1 50 |
| | (5.55) | (4.67) | (5.68) | (2.22) |
| 5 (intermittent) | 54.75 | 62.79 | 60.38 | −2.53 |
| | (4.81) | (2.92) | (3.67) | (3.13) |
| F-test | P = .000 | P = .008 | P = .000 | P = 034 |
| Total | 50 90 | 61.23 | 59.13 | −2.20 |
| | (9 09) | (3.97) | (4 19) | (3.23) |

Note.—Numbers are mean ages. SDs are in parentheses

retirement transition. On average, men retired earlier than women by about a year. Note that women's retirement age is a lot more homogeneous than men's, which is also true for the other two measures. This is in contrast to the fact that women are more varied in terms of the career paths they have traversed. In other words, women seem to be more likely to abide by retirement age norms despite their varied career experiences. Second, the trend of increasingly early retirement is clear: Later cohorts began planning earlier, expected to retire earlier, and did retire earlier. Third, between the pathway types, those following the orderly path were

the earliest in planning, closely followed by those of the high-geared type. Their predictable career pathways seem to have allowed them the room to plan ahead, which the others did not enjoy. The orderly pathway types retired the earliest, while high-geared and intermittent types retired relatively late.[15]

There is, however, a substantial discrepancy between actual and target ages of retirement: respondents, on average, retired about two years earlier than they had expected to, which is consistent with other findings (see Nestel 1985; Barfield and Morgan 1969). This holds true across gender. As expected, the latest cohort shows the largest discrepancy, more than twice that of the other two. Those of the orderly type have the largest discrepancy between actual and target retirement ages, while high-geared and steady part-time pathway types show the smallest. This is perhaps the divide between the old and new types of careers (Kanter 1989; Kotter 1995).

*Multivariate analysis.*—First, we expect gender to play an important role in the planning, expectation, and reality of scheduling retirement, for the gendered nature of life course should also be noticeable in the transition into retirement. Second, norms governing retirement as well as the structural environment around it—including corporate restructuring and introduction of early retirement incentives—have been changing over time. We expect to find some evidence of this in cohort differences among respondents. Finally, we expect career pathway type, which effectively summarizes the biographical pacing of the whole of their working lives, to influence various aspects of the retirement process. This is a straightforward life course hypothesis. A set of dummy-coded variables was used, with the intermittent type as the reference category. In addition, we control for three variables. Educational level is included as a basic control for general human capital. Tenure at the firm from which they retired and from which they receive pension is included as well. This is presumably quite important, for women tend to have shorter tenures, which might affect their retirement timing decision. We also take into account the selection bias discussed earlier.

Table 4 shows that, overall, the results are clear and consistent across models and measures. With regard to cohort—our measure of historical context—the latest cohort shows significantly earlier timing in all three

---

[15] Steady part-time pathway type shows an interesting gender-pathway type interaction. Women following this path are outliers in that they retire significantly earlier than the others ($\cong$ age 57) On the other hand, however, the men in this type are the ones who retire the latest (= age 64). Whereas men expected to retire late and actually did so, the few women in this type did the opposite Note, however, that this is the road taken by only a very small number of respondents, seven women and three men.

TABLE 4

TIMING RETIREMENT: OLS ESTIMATES

| | DEPENDENT VARIABLES | | | | | |
|---|---|---|---|---|---|---|
| | Age Began Planning | | Age Expected to Retire | | Age Retired | |
| INDEPENDENT VARIABLES | 1 | 2 | 1 | 2 | 1 | 2 |
| Gender[a] (men = 1) | −3.536** (1.165) | 6 385 (9.374) | −.199 (.412) | 5.712* (2.366) | −.110 (.352) | 6.090** (2.032) |
| Cohort:[b] | | | | | | |
| 1 (1923–28) | 3.588** (1 141) | 3.642** (1.150) | 2.672*** (.410) | 2.652*** (4.08) | 2.792*** (.348) | 2.744*** (.345) |
| 3 (1935–43) | −3.875* (1 750) | −3 742* (1.757) | −2.590*** (.635) | −2.524*** (.632) | −3 838*** (.546) | −3 758*** (.541) |
| Education (years) | −.309 (.206) | −.296 (.207) | −.065 (.074) | −.065 (.074) | .023 (.062) | .016 (.062) |
| Tenure (years) | .077 (.058) | .081 (.058) | .020 (.020) | .023 (.020) | .028 (.017) | 031 (.017) |
| Pathway type:[c] | | | | | | |
| 1 (delayed entry) | 1.534 (2.848) | 1.496 (2.853) | −1 557 (.959) | −1.565 (.953) | −1.349 (.791) | −1.350 (.783) |
| 2 (orderly) | −2.535 (2.770) | −2.345 (2.881) | −1.701 (.934) | −1.732 (.962) | −1.894* (.772) | −1.694* (.793) |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3 (high-geared) | -1.142 | -1.233 | -.959 | -.783 | -.605 | -.642 |
| | (2.666) | (2.767) | (.886) | (.913) | (.731) | (.752) |
| 4 (steady part-time) | 3.997 | 1.963 | -2.807* | -4.601** | -1.995 | -3.883** |
| | (4.460) | (4.856) | (1.369) | (1.533) | (1.159) | (1.298) |
| Gender × pathway type 2 | | -10.339 | | -5.946* | | -6.628** |
| | | (9.534) | | (2.443) | | (2.098) |
| Gender × pathway type 3 | | -9.868 | | -6.224* | | -6.155** |
| | | (9.539) | | (2.428) | | (2.085) |
| Selection bias (λ) | 1.494 | 1.420 | 1.616 | 1.545 | -1.508 | -1.533 |
| | (2.617) | (2.625) | (.961) | (.957) | (.821) | (.814) |
| Constant | 55.582 | 55.366 | 61.952 | 61.956 | 59.750 | 59.819 |
| $R^2$ | .204 | .207 | .231 | .245 | .494 | .507 |
| Adjusted $R^2$ | .175 | .173 | .210 | .220 | .480 | .491 |
| F-ratio | 7.139*** | 6.030*** | 10.983*** | 9.815*** | 36.679*** | 32.048*** |
| df | (10/279) | (12/277) | (10/365) | (12/363) | (10/376) | (12/374) |

Note.—SEs are in parentheses.

[a] Women is the omitted category

[b] Cohort 2 (1929–34) is the omitted category

[c] Pathway type 5 (intermittent career) is the omitted category.

* $P \leq .05$

** $P \leq .01$.

*** $P \leq .001$.

measures.[16] Holding other variables in the model constant, respondents in the latest cohort began to plan earlier; when they did, they expected to retire earlier; and, they did in fact retire earlier. The opposite is true of the earliest cohort. This might be due to shifting norms regarding retirement, or to the structural changes occurring in the workplace, or both, given that the former tends to reflect the latter. Note the substantial increase in the standard errors between cohorts 1 and 3 across all three measures. The norm, at least in a statistical sense, is clearly becoming looser and more diffuse over time.

Gender is significant in explaining when respondents began to plan for retirement, with men beginning to plan significantly earlier than women—by about three to four years, net of other variables. Career pathway type does not seem to discriminate significantly the timing of retirement planning. This suggests that planning may be more influenced by societal and group norms and constraints, rather than life course experience. However, it seems, albeit weakly, that those whose career pathways have been relatively smooth—the orderly and high-geared types—are more likely to begin planning early.

For both target and actual retirement ages, gender difference is contingent on career pathway type, and vice versa. On average, men are likely to retire earlier than women, and so are those following orderly and steady part-time career paths than the other three. In order to account for the contingent relationship between gender and work experience, we introduced two interaction terms between pathway type and gender into the model. We found the most substantial difference with regard to both target and actual retirement ages in men following the orderly career pathway type. They are most likely to expect to, and actually do, retire earlier. These men in traditional career paths seem to be key in the trending down of retirement age.

The omitted category, the intermittent career pathway type, was later than the others in both target and actual retirement ages, net of other variables in the model. This contrasts sharply with those experiencing delayed entry, another exclusively female pathway type. The main difference between the two is that women in the delayed entry pathway type maintained a rather stable career once they (re)entered the labor force, whereas women in the intermittent pathway type experienced a high degree of instability throughout, which seemed to affect their retirement timing as well.

Note that the model does a lot better in explaining actual timing of retirement (adjusted $R^2$ = .480 and .491) than target timing (adjusted $R^2$

[16] Given the truncated nature of our sample design, one might expect this to be the case for the actual age at retirement but not necessarily for the other two.

= .210 and .220) or timing of planning (adjusted $R^2$ = .175 and .173). The age respondents began planning is the least standardized of all three, as shown in figure 3 and table 3. In terms of target and actual retirement timing, target timing seems to be relatively more free from the constraints imposed by the factors we considered (cf. Barfield and Morgan 1969). In sum, we find all three of the factors we hypothesized—historical contexts, social heterogeneity, and biographical pacing—play significant but different roles in patterning the timing of various aspects of the retirement transition.

## To Take or Not to Take: Early Retirement Incentive

One of the changes in recent years that has considerably affected the timing of retirement is the introduction of ERIs, which are special benefit packages, mainly financial, offered to employees to encourage their retirement (often in the context of massive downsizing). To take an ERI is to leave one's job—if not the labor force—earlier than one might have done without it. Indeed, on average, those in our sample who were offered and took an ERI retired at an earlier age than those who did not by about 3 years (57.5 vs. 60.1; $F_{(1,257)}$ = 25.041; $P$ = .000). We use the model developed in the preceding section to investigate how it operates.

Table 5 reports maximum-likelihood estimation of logistic regression models. Cohort, dummy coded by two binary variables, is the only variable that matters among the three factors we modeled. Holding other variables constant and correcting for selection bias, later cohorts are significantly more likely to have taken an ERI. This also reflects their having higher rates of being offered ERIs by the employers. The effect holds true for both men and women and for all pathway types. Much like a massive shock, the downsizing/restructuring of the late 1980s and early 1990s overrides the other two clocks by fiat. Consider figure 4, where the number of retirees is plotted by year they retired. Those who took an ERI (light gray) are stacked on top of those who did not (dark gray). The steady increase in the bottom area reflects what one might expect, that is, as respondents get older, more and more of them eventually reach the point where they decide to retire. But the top area does not follow such a gradual pattern, showing instead a sharp and abrupt pike for the period roughly between 1989 and 1992. This was a period of widespread adoption of ERI, when corporate restructuring and downsizing was at its peak in the companies in our sample as in others.

We reestimated the model of retirement age presented in table 4 separately for those who took an ERI and those who did not. The contrast between the two is striking. For those who did not take an ERI, the overall results are quite similar to what we reported in table 4. For those who

# TABLE 5

## Taking an Early Retirement Incentive (ERI): Logistic Regression Estimates

| INDEPENDENT VARIABLES | Took an ERI? (Yes = 1) | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Gender[a] | .597 | .570 | −4 338 |
| | ( 317) | (.365) | (12.733) |
| Cohort:[b] | | | |
|   1 (1923–28) | −1 346* | −1 395* | −1.382* |
| | (.366) | (.374) | (.375) |
|   3 (1935–43) | 2.502* | 2.613* | 2 589* |
| | (.603) | (.651) | ( 649) |
| Education (years) | .061 | .048 | .050 |
| | ( 060) | (.063) | (.063) |
| Tenure (years) | −.013 | −.011 | −.013 |
| | (.016) | ( 019) | (.019) |
| Pathway type:[c] | | | |
|   1 (delayed entry) | | − 034 | −.032 |
| | | (.830) | (.829) |
|   2 (orderly) | | −.121 | −.226 |
| | | (.802) | (.846) |
|   3 (high-geared) | | .046 | .100 |
| | | (.760) | (.794) |
|   4 (steady part-time) | | −.635 | −.027 |
| | | (1 290) | (1.324) |
| Gender × pathway type 2 | | | 5.092 |
| | | | (12.744) |
| Gender × pathway type 3 | | | 4.839 |
| | | | (12.742) |
| Selection bias (λ) | −3 990* | −4.370* | −4 349* |
| | (.938) | (1.005) | (1.000) |
| Constant | .104 | .495 | .483 |
| −2LL | 275.26 | 267.96 | 266 74 |
| $L^2$ | 46.59 | 47.75 | 48 97 |
| $\chi^2$ (df) | 250 (6) | 240 (10) | 238 (12) |
| p-value | 001 | .001 | .001 |

Note.—SEs are in parentheses
[a] Women is the omitted category.
[b] Cohort 2 (1929–34) is the omitted category.
[c] Pathway type 5 (intermittent career) is the omitted category
* $P < 001$
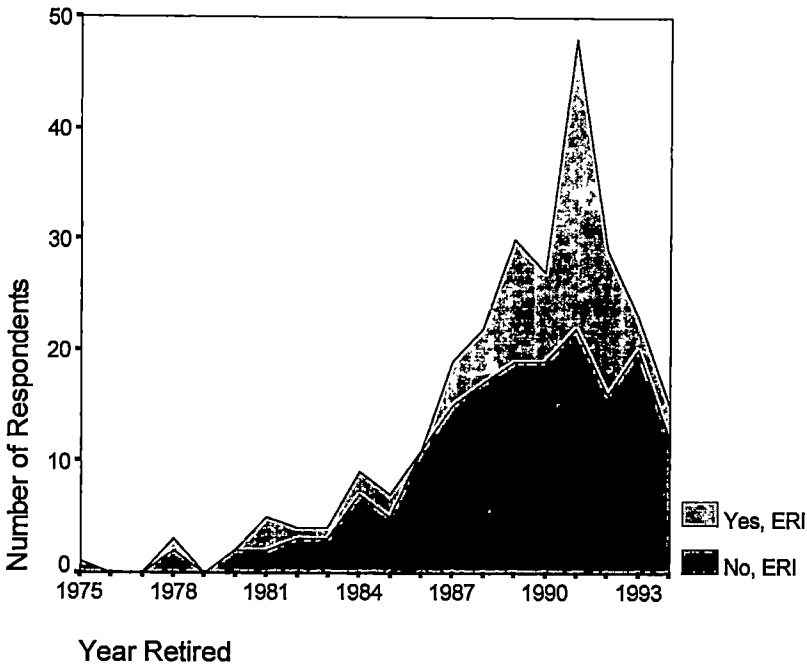
FIG. 4.—Number of respondents retired by year (CRWB, wave 1 [1994–95])
$N = 459$.

did, however, nothing but cohort turns out to be significant. Table 6 shows
the age distribution for each cohort in 1989–92. Most of those who took
an ERI come from cohorts 2 and 3, for the majority of cohort 1 had already
retired by the time ERIs began to be offered widely. Among those who
did take an ERI, it is either a matter of being the right age at the right
time, if they took advantage of the changes in the opportunity structure,
or being the wrong age at the wrong time, if they entered the transition
prematurely because of the conditions they regarded as external and coer-
cive.[17]

Postretirement (Re)Employment

Another part of the puzzle around the retirement transition is whether,
how, and why retirees go back to paid work following retirement. Our

[17] One respondent boasted that he was going to retire anyway, viewing the ERI as a
welcome windfall. Another felt it was either take the ERI or else be laid off, and was
sorry to have to leave his job.

## TABLE 6

### AGE BY COHORT, 1989

| | COHORT 1 | | | COHORT 2 | | | COHORT 3 | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Earliest | Median | Latest | Earliest | Median | Latest | Earliest | Median | Latest |
| Year born . .. . .... | 1923 | 1926 | 1928 | 1929 | 1931 | 1934 | 1935 | 1937 | 1943 |
| Age in 1989 ..... .. .. | 66 | 63 | 61 | 60 | 58 | 55 | 54 | 52 | 46 |
| Age in 1992 ... ... ... | 69 | 66 | 64 | 63 | 61 | 58 | 57 | 55 | 49 |

particular focus is on whether and how postretirement employment varies along, and is affected by, the temporal dimensions modeled to account for retirement timing. Table 7 shows the maximum-likelihood estimates of logistic regression models on the likelihood of ever taking on paid work following retirement. The first model is composed of the set of baseline variables used earlier. Unlike the results reported on the "timing" variables, human capital (i.e., education) has a large influence on postretirement employment. Specifically, respondents with higher levels of schooling are more likely to be reemployed. The odds for those with a college degree (16 years of schooling) going back to work following retirement is about twice ($1.723 = e^{16* .136}/e^{12* .136}$) that for those with only a high school degree (12 years). The delayed entry pathway type shows a significant negative effect, indicating that they are the least likely to reenter the workforce. Their relatively short, but stable, career seems to have a rather clear-cut ending. This is in sharp contrast with those in the intermittent type, another path followed exclusively by women. For this group of women, whose pathways are the most disorderly, retirement serves the demarcation function rather poorly, in the sense that they are highly likely to take on paid work again following retirement. They are more than three times ($3.300 = 1/e^{-1.194}$) more likely to do so than the delayed entry type, holding other variables constant. Given their career history, however, this postretirement employment might be just an extension of what has been an unstable work pattern.

In the second model, we add three variables to control for the effects of age and timing of interview. One might expect, for example, that those who retired in their seventies would be less likely to go back to work than those who retired in their fifties. It is also expected that those who have been in retirement longer would have been exposed to more opportunity to work again than recent retirees up to a point. The latter, which we specified as quadratic, was significant. In model 3, we introduce other variables that bear on the immediate circumstances around the retirement transition, the respondents' financial situation, and health condition at the time they retired. We expected that those who were well-prepared financially would be less likely to return to paid work, as would those who reported health problems as an important reason for retiring. We found strong support for the financial preparedness effect. The importance of poor health at the time of retirement in affecting postretirement employment differs by gender.[18] Among women, the odds of going back to work do not differ by whether or not health was a very important reason for retiring (odds ratio = .980). Among men, however, the likeli-

---

[18] Other gender interaction terms were examined but did not yield any significant result or change the overall pattern.

TABLE 7

POSTRETIREMENT EMPLOYMENT· LOGISTIC REGRESSION ESTIMATES

| INDEPENDENT VARIABLES | HAVE WORKED SINCE RETIREMENT? (Yes = 1) | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| Gender[a] (men = 1) | .022 | .116 | .770 | .881 |
| | (.280) | (.291) | ( 502) | (.520) |
| Cohort:[b] | | | | |
| 1 (≤1928) | 072 | −.270 | −.081 | − 125 |
| | (.276) | (.466) | ( 289) | (.489) |
| 3 (≥1935) | .087 | .558 | −.012 | .122 |
| | (.423) | ( 615) | (.432) | ( 636) |
| Education (years) | .136** | .138** | 149** | .142** |
| | ( 048) | (.050) | (.053) | (.054) |
| Tenure (years) | −.022 | −.022 | −.021 | −.021 |
| | (.013) | (.014) | ( 014) | (.014) |
| Pathway type:[c] | | | | |
| 1 (delayed entry) | −1.194* | −1.275* | −1.340* | −1 378* |
| | (.589) | (.604) | (.638) | (.654) |
| 2 (orderly) | −.865 | −.948 | −.984 | −1.032 |
| | (.554) | (.574) | (.606) | ( 623) |
| 3 (high-geared) | −.973 | −1.016 | −1.086 | −1.019 |
| | (.522) | (.529) | (.576) | (.583) |
| 4 (steady part-time) | −.192 | −.374 | −.219 | −.376 |
| | ( 811) | (.834) | (.850) | (.893) |
| Age at retirement | | −.000 | | −.059 |
| | | (.070) | | (.073) |
| Time since retirement | | .397** | | .314* |
| | | (.145) | | ( 154) |
| (Time since retirement)[2] | | − 019* | | −.015 |
| | | ( 009) | | (.009) |
| Finance at retirement | | | −.937*** | −.976*** |
| | | | (.270) | (.278) |
| Health at retirement | | | .020 | −.052 |
| | | | (.187) | (.193) |
| Health at retirement × gender | | | −.554 | −.560 |
| | | | ( 305) | (.314) |
| Selection bias (λ) | .394 | .076 | .260 | −.051 |
| | (.629) | (.655) | (.656) | (.688) |
| Constant | −1 790 | −2.875 | −1.133 | 1 598 |
| −2LL | 428.59 | 413.17 | 398.26 | 383.32 |
| L[2] | 21 87 | 37.29 | 39.31 | 54.25 |
| df | (376/10) | (373/13) | (362/13) | (359/16) |
| p-value | 016 | .000 | .000 | .000 |

NOTE.—SEs are in parentheses.
[a] Women is the omitted category.
[b] Cohort 2 (1929–34) is the omitted category.
[c] Pathway type 5 (intermittent career) is the omitted category.
   * P < 05.
   ** P < 01.
   *** P < 001

hood of postretirement employment of those who retired because of poor health was somewhat lower than that of those who retired for other reasons (odds ratio = .790). In sum, the circumstances of retirement seem to strongly affect the likelihood of postretirement employment.

The last model, which incorporates all the variables examined thus far, shows a consistent pattern. By and large, as expected, the factors that most strongly influence the likelihood of postretirement employment are level of human capital (i.e., education) and financial situation. There might be two sides to this dynamic, supply and demand. On the demand side is the fact that the more educated retirees are given more chances to work.[19] On the supply side, retirees who are not adequately prepared financially need work. Put together, they produce a substantial effect. For instance, other things being equal, the odds of postretirement employment for those with a college degree, who have retired without adequate financial preparation, is estimated to be about four to five times that of those with a high school degree and adequate financial preparation.

We do not find significant differences between cohorts. Gender shows only a marginal effect, mainly due to the differential responses to the health situation at retirement. Pathway type—the way in which one's career came to an end, in particular—also has a marginally significant effect, showing intriguing differences. The delayed entry career pathway type clearly stands out from the pack. Less clear, but still substantial, are the coefficients for the orderly and high-geared pathway types (.05 < P < .10). Their odds ratios are .356 and .361, respectively. In other words, net of the other variables in the model, only a few of those following these two pathways are likely to return to work once they retire. Steady part-time and intermittent career pathway types are the ones with the most blurred boundary between work and nonwork. By and large, the effects of the three dimensions were not as strong as the ones we have found for retirement timing behaviors.

## DISCUSSION AND CONCLUSIONS

We proposed a life course model of retirement in which retirement is timed on three dimensions—(historical) context, (social) heterogeneity, and (biographical) pacing. Using life history data for retired men and women, we found that all three affect various aspects of retirement and their temporal patterning, yet in different ways and for different reasons. The results of the analysis are hence rather complex. They can, however, be integrated

---

[19] Another possibility is that the more educated might have a greater "taste" for work (Parnes and Less 1985).

into and understood within a larger framework of our proposed model of multiplex time.

First, historical context, measured by cohort, is significant in explaining differences in the age respondents began planning, the target retirement age, the actual retirement age, and the likelihood of their taking an ERI. The effect on the latter, moreover, is distinguishable from the effect of the long-term trending down of retirement age. Second, we find gender differences in retirement timing; on average, men are more likely to be earlier than women on three of the variables associated with the temporal patterning of the retirement process. But the effects of gender on these variables are largely contingent upon the different career paths men and women have traveled, that is, the gendered nature of career pathways. Third, and finally, career pathway type, which summarizes the biographical pacing embodied in respondents' work history, shows the strongest effect on the actual timing of retirement. On average, those following an orderly career path retired the earliest, and within it, men more so than women. Career pathway type also affects the expectation of retirement timing but not its planning. We find no significant differences between men and women or between pathway types in terms of likelihood of taking an ERI, which strongly suggests that it might be period specific, experienced by all workers of certain ages as an exogenous shock. Among the dependent variables we examined, postretirement employment is distinct from the others in that it is the least affected by the three factors we hypothesized. Pathway type is the only significant, albeit weak, discriminating factor. Postretirement employment has long been considered an integral part of the overall process of the retirement transition, but it apparently is affected by something other than these three factors.

These findings serve to illuminate the puzzle with which we began: the peculiar pairing of the lowering of average age at retirement and its increasing variability. Context, heterogeneity, and pacing are all related to the increasingly early retirement age. The context effect, measured by cohort, shows a strong and consistent trend. Historical, structural, and institutional contexts in the latter half of the 20th century have steadily become more conducive to early retirement. Men, who have historically made up the majority of the labor force, are more likely to retire earlier than women, especially those following traditional career paths. In short, the decline in average retirement age seems to have been largely driven by long-term historical trends, affecting the traditional workforce most strongly.

On the other hand, all three factors have played parts, either alone or in combination, in making retirement timing more heterogeneous. Histori-

cal and institutional contexts have become less and less effective in enforcing any one age norm, primarily due to the developments that were in conflict with one another. Occasional external shocks, like the recent spell of massive and widespread downsizing, also produce more variability in timing. Given the gender difference in retirement timing, mainly due to the gap between men's and women's work experiences, the increasing rate of women's labor force participation has also contributed to increasing variability. Finally, the different pacing of retirement transition by those who experience various career pathways remains a major source of variability, partly due to the parallel development of increasing individuation and decreasing social regulation per se and partly due to changes occurring in the occupational and industrial structures that produce a more heterogeneous mix of career trajectories.

In sum, the temporal regularity with regard to retirement in the life course regime of the middle of the 20th century was a product of the convergence of diverse institutional features combined with the centrality of one group of the population, a large core of men on the traditional career paths—mainly the orderly pathway type—anchoring the norms. Recent changes in retirement timing might not have happened if at least one of these circumstances had held firm. But both have undergone changes, and the link between the two has become unhinged, producing the current situation.

The model we proposed seems to adequately explain changes in retirement timing and can be an effective framework for the study of life course institutionalization and deinstitutionalization in general, with a few qualifications. First, we need to be cautious in interpreting and generalizing these findings, given our data limitations. While our respondents are drawn randomly from six companies in upstate New York, they are not a representative national sample. For example, the women represented in our sample are rather a very select group, especially for the cohorts we studied. Whether the five career pathway types we identified comprehensively represent the population at large is, for instance, a question to be answered in the future with appropriate data.[20] The retirement process itself is also sensitive to historical context and continues to be in flux. The data also do not permit much variance in—and hence, the model did not include—other aspects of population heterogeneity, such as race or indus-

---

[20] See D'Unger et al. (1998). Most of the studies employing the sequence analysis technique so far have been based on relatively delimited data settings, e.g., German musicians in the 18th century (Abbott and Hrycak 1990), employees at Lloyds Bank (Stovel et al. 1996), and female senior executives in finance (Blair-Loy 1999).

try, which also could be pertinent (Henretta 1992; Rosenfeld 1980; Barfield and Morgan 1969).[21] Nor do our data provide a sufficiently wide cohort/period range to fully address the question of structural changes and their effects.

Second, in terms of the model, we need to elaborate and extend it further. For example, although tenure might serve as a rough proxy for eligibility and amount of pension, the model does not explicitly take into account the financial aspects of retirement timing. There is a large body of literature that emphasizes the role economic incentives play in retirement decisions (e.g., Hardy et al. 1996; Barfield and Morgan 1969). The findings to date on the link between the early exit trend and the financial incentives in retirement systems, however, are rather mixed (see Guillemard and Rein [1993] for a review). Some argue to have established a direct cause-effect relationship, while others argue that the link is far from certain (Doeringer 1990). Yet the baseline model we have used and the results we obtained seem to be robust and consistent with the finding that early exit may well have been imposed rather than chosen (Guillemard and Rein 1993, p. 479).

Also, the model focused on experiences only in work domain, yet events within single life domain usually cannot be fully understood without reference to events in other life domains (Settersten and Mayer 1997; Mayer and Tuma 1990). The most crucial is, of course, the family, where the notion of linked or interdependent lives becomes particularly salient (Elder 1994). For example, Henretta et al. (1993) found that a woman's employment during the childrearing years is associated with her earlier retirement, especially following her husband's retirement (also see Henretta and O'Rand 1983; O'Rand and Henretta 1982). Instead of looking at individual careers, hence, we need to be able to consider "coupled" careers as well as the interplay between family and work "careers" (Han and Moen 1999).

The last two decades have seen a growing interest and an explosive increase in theoretical and empirical studies of the social organization of the life course (see Mayer and Schoepflin [1989] for a review). Although a variety of relevant factors have been discussed, the multiplex nature of temporality has eluded an integrative model that can explain individual life events and social patterns of life trajectories within a common conceptual and empirical framework and represent the social processes that generate these events and trajectories (Mayer and Tuma 1990, p. 5). The conceptual model we presented is certainly one that takes into account multiple components, multiple levels, and multiple times. Recent changes

[21] For instance, the largest minority group consists of 22 African-Americans, accounting for only 2.9% of the initial sample

in retirement timing provided a research site to illustrate the utility of such a model.

Finally, while transitions and trajectories are both crucial in understanding temporal patterning of life course, the analysis of transitions has been far more developed than that of trajectories. Much of the reason for this appears to be methodological, with increased attention to transitions fueled by growing interest and availability of event history analysis techniques (Pavalko 1995). As a result, the issue of timing has been typically addressed in an event history framework, where the focus is on the rate at which some event or transition, such as retirement, occurs. Noting the imbalance, we focused on trajectories, introducing a novel methodological technique, sequence analysis (Abbott 1995a), and empirically assessing career paths leading to retirement. By focusing on the overall career pathway, and embedding the retirement transition within it, we were able to look into the temporal dimensions underlying various retirement-related behaviors. We believe the two key concepts—transitions and trajectories—and the two methodological frameworks—event history analysis and sequence analysis—can and should be brought together more closely.

APPENDIX

Obtaining Interelement Distances

Interelement distances are obtained from the analysis of the complete transition matrices that report the distribution of transitions for all retirees in our sample over the entire period, from age 30 to retirement, for each of the three dimensions. For example, table A3 shows the number of transitions between the five work status categories. We first symmetrized and normalized the matrix. We then computed interelement distances based on comparisons of rows and columns using Euclidean distance algorithm, producing the results in table A4. The procedure was repeated separately for occupation and organization. Further technical details of the procedure used in the analysis, including the interelement distance matrices for occupation and organization, are available upon request.

APPENDIX

## TABLE A1

### MODELING SELECTION BIAS ($\lambda$) AND DESCRIPTIVE STATISTICS

| | PROBIT REGRESSION COEFFICIENT | OVERALL MEAN | GROUP MEANS | | |
| VARIABLE | | | Active | Retired | P |
|---|---|---|---|---|---|
| Gender (men = 1) ............ ...... | −.163 | .504 | .461 | .533 | |
| | ( 144) | | | | |
| Cohort | | | | | |
| 1 (1924–28) ... .................... .. | .593** | 244 | .049 | .373 | *** |
| | (.189) | | | | |
| 2 (1929–34) ............. ........... | · · · | .299 | 122 | .417 | *** |
| 3 (1935–43) ............ . ........ ... | −1.640*** | 457 | 829 | .210 | *** |
| | (.149) | | | | |
| Education   .. .. .... . ............... | −.058* | 13.881 | 14.310 | 13.582 | *** |
| | (.023) | | | | |
| Tenure ...................... . . .. ....... | .027*** | 19.527 | 16.468 | 21 632 | *** |
| | ( 006) | | | | |
| Company | | | | | |
| B ... ...................... .. .. ... .. .. | −2.679*** | 140 | .194 | .105 | *** |
| | (.452) | | | | |
| C . ... .. .. . .. ... .............. .. | −2.299*** | .270 | 220 | .304 | * |
| | (.435) | | | | |
| K ............. ............ .. ........... | −1.691*** | .223 | .240 | .212 | |
| | (.428) | | | | |
| N   .. ...... .. . .. ... .................... | −2 080*** | .142 | .165 | .127 | |
| | (.443) | | | | |
| S ...... ...... ... .... . . .. ............... | −2 761*** | 106 | .181 | .057 | *** |
| | (.463) | | | | |
| X ................................ ........ | · · | .118 | .000 | .197 | *** |
| Intercept ................................. | 3 392 | | | | |
| | ( 538) | | | | |
| N ...................................... .. ....... | | 762 | 304 | 458 | |

NOTE —P in this column indicates *F*-test on the difference between the two group means  SEs are in parentheses
* $P < .05$.
** $P < .01$.
*** $P < .001$.

228

TABLE A2

OCCUPATIONAL CATEGORY CODES

| CRWB Code | Description | 1990 Census Three-Digit Codes | SEI Score (S*) |
|---|---|---|---|
| 01 | Executives, administrative, and managerial | 003–022 | 66 80 |
| 02 | Management related occupations | 023–037 | 69.46 |
| 03 | Engineers, architects, and surveyors | 043–063 | 85.80 |
| 04 | Mathematical and computer scientists | 064–068 | 83.46 |
| 05 | Natural scientists | 069–083 | 83.41 |
| 06 | Health, diagnosing | 084–089 | 95.29 |
| 07 | Health, assessment and treating | 095–106 | 71.95 |
| 08 | Teachers, post-secondary | 113–154 | 82.84 |
| 09 | Teachers, except post-secondary | 155–159 | 74.69 |
| 10 | Counselors, educational and vocational | 163 | 81.00 |
| 11 | Librarians, archivists, and curators | 164–165 | 71.85 |
| 12 | Social scientists and urban planners | 166–173 | 83.93 |
| 13 | Social, recreation, and religious workers | 174–177 | 68.27 |
| 14 | Lawyers and judges | 178–179 | 91.84 |
| 15 | Writers, artists, entertainers, athletes, and designers | 183–199 | 65.19 |
| 16 | Health, technologists, and technicians | 203–208 | 53.15 |
| 17 | Technologists and technicians, except health | 213–218 | 61.22 |
| 18 | Science technicians | 223–225 | 59.60 |
| 19 | Technicians, except health, engineering, and science | 226–235 | 67.48 |
| 20 | Sales, supervisors and proprietors | 243 | 51.00 |
| 21 | Sales representatives, finance and business services | 253–257 | 66.31 |
| 22 | Sales representatives, commodities, except retail | 258–259 | 65.33 |
| 23 | Sales workers, retail and personal services | 263–278 | 36 58 |
| 24 | Sales related | 283–285 | 44.00 |
| 25 | Administrative support, supervisors | 303–307 | 53.58 |
| 26 | Computer equipment operators | 308–309 | 46.18 |
| 27 | Secretaries, stenographers, and typists | 313–315 | 38 26 |
| 28 | Information clerks | 316–323 | 40.00 |
| 29 | Records processing, except financial | 325–336 | 43.80 |
| 30 | Financial records processing | 337–344 | 37.36 |
| 31 | Duplicating, mail and other office machine operators | 345–347 | 35.74 |
| 32 | Communications equipment operators | 348–353 | 31.40 |
| 33 | Mail and message distributing | 354–357 | 48.17 |

TABLE A2 (Continued)

| CRWB Code | Description | 1990 Census Three-Digit Codes | SEI Score (S*) |
|---|---|---|---|
| 34 ............. | Material recording, scheduling, and distributing clerks | 359–374 | 37.57 |
| 35 ... .. .. ... .... | Adjusters and investigators | 375–378 | 52.58 |
| 36 .............. | Miscellaneous administrative support | 379–389 | 38.94 |
| 37 ............ | Private household | 403–407 | 25.03 |
| 38 ............ | Protctive service, supervisors | 413–415 | 62.55 |
| 39 ... .. .. .. ... | Firefighting and fire prevention | 416–417 | 52 16 |
| 40 ............ | Police and detectives | 418–424 | 63.00 |
| 41 ............ | Guards | 425–427 | 42.81 |
| 42 ... ... .. .. . .. | Food preparation and service | 433–444 | 31.33 |
| 43 ... .. .. .. . . . | Health service | 445–447 | 30.70 |
| 44 ............ . . | Cleaning and building service, except household | 448–455 | 27.08 |
| 45 . ... ... ... ... | Personal service | 456–469 | 31.96 |
| 46 ... .. . . ... ... | Farm operators and managers | 473–476 | 37.88 |
| 47 ............ | Farm occupations, except managerial | 477–484 | 27.38 |
| 48 ... ... .. .. ... | Related agricultural | 485–489 | 31.90 |
| 49 ... . .. ... ..... | Forestry and logging | 494–496 | 31.90 |
| 50 ... . .. ... ..... | Fishers, hunters, and trappers | 497–499 | 37.50 |
| 51 ............... | Mechanics and repairers, supervisors | 503 | 49.00 |
| 52 . ... ... ... ... | Mechanics and repairers | 505–533 | 38 27 |
| 53 ............. | Heating, air conditioning, and refrigeration mechanics | 534 | 39.00 |
| 54 ............ | Miscellaneous mechanics and repairers | 535–549 | 40.45 |
| 55 ... . . .. ... ... | Construction trades, supervisors | 553–558 | 49 08 |
| 56 ... .. . . .. .. .. | Construction trades | 563–599 | 35 51 |
| 57 ... . . . ... .. .. | Extractive occupations | 613–617 | 39.53 |
| 58 ... ............ | Precision production, supervisors | 628 | 49.00 |
| 59 . ... ... ... ... | Precision production, except supervisors | 634–699 | 34.66 |
| 60 . ... .. ... ..... | Machine operators, assemblers, and inspectors | 703–799 | 27.36 |
| 61 ... .. .. .. ... .. | Transportation and material moving | 803–859 | 33.33 |
| 62 ............ | Handlers, equipment cleaners, helpers, and laborers | 864–889 | 30.09 |
| 63 ... .. .. .. ... .. | Military | 903–905 | 55.00 |
| 64 ... ... ... ..... | Not employed | 909 | (n/a) |

Note.—SEI score ($S^*$) is assigned to each occupation code used in the Cornell Retirement and Well-Being study (CRWB) as follows $S_i^* = \sum w_j S_j$, where $i$ indexes CRWB occupation code, and $j$ indexes census three-digit code $w_j$ denotes $j$'s relative frequency within $i$—$f_j/N$, —taken from the GSS, and $S_j$ is from Nakao and Treas (1994)

TABLE A3

TRANSITIONS BETWEEN WORK STATUS CATEGORIES

| Element Code | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 Full-time ........... . . ......... ........ .. . ...... ......... | 991 | 2 | 25 | 0 | 66 |
| 2 Alternate, F/T, P/T ... ...... . . ................... ..... ..... | 6 | 0 | 1 | 0 | 2 |
| 3 Part-time ...... ... .. .. ...... ..... ... .......................... ... . | 68 | 1 | 15 | 0 | 17 |
| 4 Sporadic, Seasonal ............. . ........................... .. | 4 | 0 | 1 | 0 | 1 |
| 5 Unemployed/OLF ............... .. .... . .. ................. | 153 | 5 | 44 | 5 | 0 |

TABLE A4

INTERELEMENT DISTANCES

| Element Code | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 Full-time ................. .., ... ......... ...... .... .. .. ...... | .000 | 793 | .782 | .861 | .682 |
| 2 Alternate, F/T, P/T .. ... .. ....... ......... . .. .. ...... | .793 | .000 | .502 | 1 000 | .223 |
| 3 Part-time .. .. ... ....................... ................... . . . . | 782 | .502 | .000 | .680 | .313 |
| 4 Sporadic, Seasonal .. ............... ...... _ ................. | 861 | 1.000 | .680 | .000 | .180 |
| 5 Unemployed/OLF ...... ... ...... ... ......... .. .. .. .. .. .. | .682 | .223 | .313 | .180 | .000 |

REFERENCES

Abbott, Andrew. 1995a. "Sequence Analysis: New Methods for Old Ideas." *Annual Review of Sociology* 21:93–113.
———. 1995b. OPTIMIZE, ver. 2.12. University of Chicago, Department of Sociology.
Abbott, Andrew, and Emily Barman. 1997. "Sequence Comparison via Alignment and Gibbs Sampling: A Formal Analysis of the Emergence of the Modern Sociological Article." *Sociological Methodology* 27:47–87
Abbott, Andrew, and Alexandra Hrycak. 1990. "Measuring Resemblance in Sequence Data: An Optimal Matching Analysis of Musicians' Careers." *American Journal of Sociology* 96:144–85.
Anderson, Michael 1985. "The Emergence of the Modern Life Cycle in Britain." *Social History* 10:69–87.
Atchley, Robert C. 1982. "Retirement as a Social Institution." *Annual Review of Sociology* 8:263–87.
Barfield, Richard E., and James N. Morgan. 1969 *Early Retirement: The Decision and the Experience.* Ann Arbor: University of Michigan, Institute for Social Research.
Barley, Stephen R. 1989. "Careers, Identities, and Institutions: The Legacy of the Chicago School of Sociology." Pp. 41–65 in *Handbook of Career Theory,* edited by Michael B. Arthur, Douglas T. Hall, and Barbara S Lawrence. New York: Cambridge University Press.
Becker, Gary S. [1964] 1993. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education* Chicago: University of Chicago Press.

Blair-Loy, Mary. 1999. "Career Patterns of Executive Women in Finance: An Optimal Matching Analysis." *American Journal of Sociology* 104 (5): 1346–97.

Breiger, Ronald L. 1995. "Social Structure and the Phenomenology of Attainment." *Annual Review of Sociology* 21:115–36

Burt, Ronald S. 1983. "Cohesion versus Structural Equivalence as a Basis for Network Subgroups " Pp. 262–82 in *Applied Network Analysis: A Methodological Introduction,* edited by Ronald S. Burt and Michael J. Minor. Beverly Hills, Calif.: Sage.

———— 1991. "Measuring Age as a Structural Concept." *Social Networks* 13:1–34.

Campbell, Richard T., and Angela M. O'Rand 1988. "Settings and Sequences: The Heuristics of Aging Research." Pp. 58–79 in *Emergent Theories of Aging,* edited by James E. Birren and Vern L Bengston. New York: Springer.

Chan, Tak Wing 1995. "Optimal Matching Analysis: A Methodological Note on Studying Career Mobility." *Work and Occupations* 22:467–90

Dannefer, Dale. 1988. "What's in a Name? An Account of the Neglect of Variability in the Study of Aging." Pp. 356–84 in *Emergent Theories of Aging,* edited by James E. Birren and Vern L. Bengtson. New York: Springer.

DeViney, Stanley, and Angela M O'Rand. 1988. "Gender-Cohort Succession and Retirement among Older Men and Women, 1951 to 1984." *Sociological Quarterly* 29: 525–40.

Doeringer, Peter B., ed. 1990. *Bridges to Retirement: Older Workers in a Changing Labor Market.* Ithaca, N.Y.: ILR Press.

D'Unger, Amy, Kenneth C. Land, Patricia L. McCall, and Daniel S. Nagin. 1998. "How Many Latent Classes of Delinquent/Criminal Careers? Results from Mixed Poisson Regression Analyses." *American Journal of Sociology* 103:1593–1630.

Ekerdt, David J., and Stanley DeViney. 1993. "Evidence for a Preretirement Process among Older Male Workers." *Journal of Gerontology: Social Sciences* 48:S35–S43

Elder, Glen H., Jr. 1985. "Perspectives on the Life Course." Pp. 23–49 in *Life Course Dynamics: Trajectories and Transitions, 1968–1980,* edited by Glen H. Elder, Jr. Ithaca, N.Y. Cornell University Press.

———— 1992. "The Life Course." Pp. 1120–30 in *The Encyclopedia of Sociology,* edited by Edgar F. Borgatta and Marie L. Borgatta. New York: Macmillan.

————. 1994 "Time, Human Agency, and Social Change: Perspectives on the Life Course." *Social Psychology Quarterly* 57 4–15.

————. 1995. "The Life Course Paradigm: Social Change and Individual Development." Pp. 101–39 in *Examining Lives in Context: Perspectives on the Ecology of Human Development,* edited by Phyllis Moen, Glen H. Elder, Jr., and Kurt Lüscher. Washington, D.C.: American Psychological Association.

Elder, Glen H., Jr., and Angela M. O'Rand. 1995. "Adult Lives in a Changing Society." Pp. 452–75 in *Sociological Perspectives on Social Psychology,* edited by Karen S. Cook, Gary Alan Fine, and James S. House. Boston: Allyn & Bacon.

Elder, Glen H., Jr, Eliza K. Pavalko. 1993. "Work Careers in Men's Later Years: Transitions, Trajectories, and Historical Change." *Journal of Gerontology: Social Sciences* 48:S180–S91.

Featherman, David L 1986. "Biography, Society, and History: Individual Development as a Population Process " Pp. 99–149 in *Human Development and the Life Course: Multidisciplinary Perspectives,* edited by Aage B. Sørensen, Franz E. Weinert, and Lonnie R. Sherrod. Hillsdale, N.J.: L. Erlbaum.

George, Linda K. 1993. "Sociological Perspectives on Life Transitions." *Annual Review of Sociology* 19:353–73.

Graebner, William 1980. *A History of Retirement: The Meaning and Function of an American Institution, 1885–1978.* New Haven, Conn.: Yale University Press.

Greene, William H. 1993. *Econometric Analysis,* 2d ed New York. Macmillan.

Guillemard, Anne Marie, and Herman van Gunsteren. 1991. "Pathways and Prospects: A Comparative Interpretation of the Meaning of Early Exit." Pp. 362–88 in *Time for Retirement: Comparative Studies of Early Exit from the Labor Force,* edited by Martin Kohli, Martin Rein, Anne Marie Guillemard, and Herman van Gunsteren. New York: Cambridge University Press.

Guillemard, Anne Marie, and Martin Rein. 1993. "Comparative Patterns of Retirement: Recent Trends in Developed Societies." *Annual Review of Sociology* 19:469–503.

Han, Shin-Kap, and Phyllis Moen. 1999. "Coupled Careers: Men's and Women's Pathways through Work and Marriage in the United States." In *Couples' Careers in a Dynamic Perspective,* edited by Hans-Peter Blossfeld. In press.

Hardy, Melissa A., Lawrence Hazelrigg, and Jill Quadagno. 1996 *Ending a Career in the Auto Industry: "30 and Out."* New York: Plenum.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47:153–61.

Henretta, John C. 1992. "Uniformity and Diversity: Life Course Institutionalization and Late-Life Work Exit." *Sociological Quarterly* 33:265–79.

Henretta, John C., and Hyunkee Lee. 1996. "Cohort Differences in Men's Late-Life Labor Force Participation." *Work and Occupations* 23:214–35.

Henretta, John C., and Angela M. O'Rand. 1983. "Joint Retirement in the Dual Worker Family." *Social Forces* 62:504–20

Henretta, John C., Angela M. O'Rand, and Christopher G. Chan. 1993. "Joint Role Investments and Synchronization of Retirement. A Sequential Approach to Couples' Retirement Timing." *Social Forces* 71:981–1000.

Hogan, Dennis P. 1978. "The Variable Order of Events in the Life Course." *American Sociological Review* 43:573–86.

Hogan, Dennis P., and Nan Marie Astone. 1986 "The Transition to Adulthood." *Annual Review of Sociology* 12:109–30.

Hughes, Everett C. [1937] 1994. "Institutional Office and the Person." Pp. 136–45 in *On Work, Race, and the Sociological Imagination,* edited by Lewis A. Coser. Chicago: University of Chicago Press.

Jacobs, Klaus, Martin Kohli, and Martin Rein. 1991 "Testing the Industry-Mix Hypothesis of Early Exit." Pp. 67–96 in *Time for Retirement: Comparative Studies of Early Exit from the Labor Force,* edited by Martin Kohli, Martin Rein, Anne-Marie Guillemard, and Herman van Gunsteren. New York: Cambridge University Press.

Juster, F. Thomas. 1992. Health and Retirement Survey. University of Michigan, Institute for Social Research, Survey Research Center

Kalleberg, Arne L., David Knoke, and Peter V. Marsden. 1995. "Interorganizational Networks and the Changing Employment Contract." *Connections* 18:32–49.

Kanter, Rosabeth M. 1989. *When Giants Learn to Dance.* New York: Simon & Schuster.

Kohli, Martin, and Martin Rein. 1991. "The Changing Balance of Work and Retirement." Pp. 1–35 in *Time for Retirement: Comparative Studies of Early Exit from the Labor Force,* edited by Martin Kohli, Martin Rein, Anne-Marie Guillemard, and Herman van Gunsteren. New York: Cambridge University Press.

Kotter, John P. 1995. *The New Rules: How to Succeed in Today's Post-Corporate World.* New York: Free Press.

Kreps, Juanita, and Robert Clark. 1975 *Sex, Age, and Work: The Changing Composition of the Labor Force.* Baltimore Johns Hopkins University Press.

Maddox, George L. 1987. "Aging Differently." *Gerontologist* 27:557–64.

Marini, Margaret Mooney. 1984. "Age and Sequencing Norms in the Transition to Adulthood." *Social Forces* 63:229–44.

Masnick, George, and Mary Jo Bane 1980. *The Nation's Families: 1960–1990.* Cambridge, Mass.: Joint Center for Urban Studies of MIT and Harvard University.

Mayer, Karl Ulrich. 1986. "Structural Constraints in the Life Course." *Human Development* 29:163–70.

Mayer, Karl Ulrich, and Urs Schoepflin. 1989. "The State and the Life Course." *Annual Review of Sociology* 15.187–209.

Mayer, Karl Ulrich, and Nancy B Tuma 1990 "Life Course Research and Event History Analysis: An Overview." Pp. 3–20 in *Event History Analysis in Life Course Research,* edited by Karl Ulrich Mayer and Nancy B. Tuma. Madison: University of Wisconsin Press

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings.* New York· National Bureau of Economic Research.

Mitchell, Olivia S , and Gary S. Fields. 1982. "The Effects of Pensions and Earnings on Retirement· A Review Essay." *Research in Labor Economics* 5:115–55.

Modell, John. 1989. *Into One's Own. From Youth to Adulthood in the United States, 1920–1975.* Berkeley and Los Angeles: University of California Press.

Moen, Phyllis. 1985. "Continuities and Discontinuities in Women's Labor Force Activity." Pp. 113–55 in *Life Course Dynamics: Trajectories and Transitions, 1968–1980,* edited by Glen H. Elder, Jr. Ithaca. N.Y : Cornell University Press.

———. 1996*a.* "Gender, Age, and the Life Course " Pp. 171–87 in *Handbook of Aging and the Social Sciences,* 4th ed. Edited by Robert H. Binstock and Linda K. George. San Diego, Calif.: Academic Press.

——— 1996*b.* "A Life Course Perspective on Retirement, Gender, and Well-Being." *Journal of Occupational Health Psychology* 1:131–44.

Mutchler, Jan E., Jeffrey A. Burr, Amy M. Pienta, and Michael P. Massaglil. 1997. "Pathways to Labor Force Exit: Work Transitions and Work Instability." *Journal of Gerontology: Social Sciences* 52B:S4–S12

Nakao, Keiko, and Judith Treas. 1994. "Updating Occupational Prestige and Socioeconomic Scores. How the New Measures Measure Up." *Sociological Methodology* 24:1–72.

Nestel, Gilbert. 1985. "Retirement Expectation and the Timing of Retirement." Pp. 79–89 in *Retirement among American Men,* edited by Herbert S. Parnes et al. Lexington, Mass.: Lexington Press.

Neugarten, Bernice L. 1979. "Time, Age, and the Life Cycle." *American Journal of Psychiatry* 136:887–94.

Neugarten, Bernice L., and Gunhild O Hagestad. 1976. "Age and the Life Course." Pp 25–55 in *Handbook of Aging and the Social Sciences,* edited by Robert H. Binstock and Ethel Shanas. New York: Van Nostrand Reinhold.

O'Rand, Angela M. 1996 "Structuration and Individualization· The Life Course as a Continuous, Multilevel Process " Pp. 3–16 in *Generating Social Stratification: Toward a New Research Agenda,* edited by Alan C Kerckhoff. Boulder, Colo.: Westview Press

O'Rand, Angela M., and John C. Henretta. 1982. "Delayed Career Entry, Industrial Pension Structure, and Early Retirement in a Cohort of Unmarried Women." *American Sociological Review* 47:365–73.

Osterman, Paul. 1988 *Employment Futures: Reorganization, Dislocation, and Public Policy.* New York: Oxford University Press.

Pampel, Fred C., and Sookja Park. 1986. "Cross-National Patterns and Determinants of Female Retirement." *American Journal of Sociology* 91·932–55.

Pampel, Fred C., and Jane A. Weise. 1983. "Economic Development, Pension Policies, and the Labor Force Participation of Aged Males: A Cross-National. Longitudinal Approach." *American Journal of Sociology* 89:350–72

Parnes, Herbert S., and Lawrence Less. 1985. "The Volume and Pattern of Retirements, 1966–1981." Pp. 57–77 in *Retirement among American Men*, edited by Herbert S. Parnes. Lexington, Mass.: Lexington Press.

Parnes, Herbert S., and Gilbert Nestel. 1981. "The Retirement Experience." Pp. 155–97 in *Work and Retirement: A Longitudinal Study of Men*, edited by Herbert S. Parnes. Cambridge, Mass.: MIT Press.

Pavalko, Eliza K 1995. "Transitions and Trajectories as Concepts for Understanding Social Change and Individual Lives." Paper presented at the 1995 ASA annual meeting, Washington, D.C.

Pavalko, Eliza K., Glen H. Elder, Jr., and Elizabeth C. Clipp. 1993. "Work Lives and Longevity: Insights from a Life Course Perspective." *Journal of Health and Social Behavior* 34:363–80.

Petersen, Trond, and Seymour Spilerman. 1990. "Job Quits from an Internal Labor Market." Pp. 69–95 in *Event History Analysis in Life Course Research*, Karl Ulrich Mayer and Nancy B. Tuma Madison University of Wisconsin Press.

Quinn, Joseph F., and Richard V Burkhauser. 1994. "Retirement and Labor Force Behavior of the Elderly." Pp. 50–101 in *Demography of Aging*, edited by Linda Martin and Samuel Preston. Washington, D.C.: National Academy of Science.

Quinn, Joseph F., Richard V. Burkhauser, and Daniel A. Myers. 1990. *Passing the Torch: The Influence of Financial Incentives on Work and Retirement*. Kalamazoo, Mich.: W. E. Upjohn Institute on Employment Research.

Quinn, Robert P., and Graham L. Staines. 1979. The 1977 Quality of Employment Survey. University of Michigan, Institute for Social Research, Survey Research Center.

Ransom, Roger L., and Richard Sutch. 1986. "The Labor of Older Americans: Retirement of Men On and Off the Job, 1870–1937." *Journal of Economic History* 46· 1–30.

Riley, Matilda W , and John W. Riley, Jr. 1994. "Structural Lag: Past and Future." Pp. 15–36 in *Age and Structural Lag: Society's Failure to Provide Meaningful Opportunities in Work, Family, and Leisure*, edited by Matilda W Riley, Robert L. Kahn, and Anne Foner. New York: Wiley & Sons

Rindfuss, Ronald R., C Gray Swicegood, and Rachel A. Rosenfeld. 1987. "Disorder in the Life Course: How Common and Does It Matter?" *American Sociological Review* 52:785–801.

Rosenbaum, James E. 1984. *Career Mobility in a Corporate Hierarchy*. Orlando, Fla : Academic Press.

Rosenfeld, Rachel A. 1980. "Race and Sex Differences in Career Dynamics." *American Sociological Review* 45:583–609.

———. 1992. "Job Mobility and Career Processes." *Annual Review of Sociology* 18: 39–61.

Ryder, Norman B. 1965. "The Cohort as a Concept in the Study of Social Change." *American Sociological Review* 30:843–61.

Settersten, Richard A., Jr., and Karl Ulrich Mayer. 1997. "The Measurement of Age, Age Structuring, and the Life Course." *Annual Review of Sociology* 23.233–61

Sørensen, Aage B , and Nancy B. Tuma 1981. "Labor Market Structures and Job Mobility." *Research in Social Stratification and Mobility* 1.67–94.

Spenner, Kenneth I , Luther B Otto, and Vaughn R. A Call. 1982. *Career Lines and Career*. Lexington, Mass.: Lexington Press.

Spilerman, Seymour. 1977. "Careers, Labor Market Structure, and Socioeconomic Attainment." *American Journal of Sociology* 83:551–93.

Stovel, Katherine. 1996. DISTANCE. University of North Carolina at Chapel Hill, Department of Sociology.

Stovel, Katherine, Michael Savage, and Peter Bearman. 1996 "Ascription into

Achievement: Models of Career Systems at Lloyds Bank, 1890–1970." *American Journal of Sociology* 102:358–99.

Streib, Gordon F., and Clement J. Schneider. 1971. *Retirement in American Society: Impact and Process.* Ithaca, N.Y.: Cornell University Press

Tomaskovic-Devey, Donald 1993. "Labor-Process Inequality and the Gender and Race Composition of Jobs." *Research in Social Stratification and Mobility* 12:215–47.

Tuma, Nancy B., and Michael T. Hannan. 1984. *Social Dynamics: Models and Methods.* Orlando, Fla.: Academic Press.

Tuma, Nancy B., and Gary D. Sandefur. 1988. "Trends in the Labor Force Activity of the Elderly in the United States, 1940–1980." Pp. 38–83 in *Issues in Contemporary Retirement,* edited by Rita Ricardo-Campbell and Edward P. Lazear. Stanford, Calif.: Hoover Institution Press

Waite, Linda J. 1981. "U.S. Women at Work." *Population Bulletin* 36:1–45.

White, Harrison C., Scott A. Boorman, and Ronald L. Breiger. 1976. "Social Structure from Multiple Networks. I.: Blockmodels of Roles and Positions." *American Journal of Sociology* 81:730–80.

Wilensky, Harold L. 1961. "Orderly Careers and Social Participation: The Impact of Work History on Social Integration in the Middle Class." *American Sociological Review* 26:521–39.

Winship, Christopher, and Robert D. Mare. 1992. "Models for Sample Selection Bias." *Annual Review of Sociology* 18:327–50.

# Book Reviews

*The Struggle for Water: Politics, Rationality, and Identity in the American Southwest.* By Wendy Nelson Espeland. Chicago: University of Chicago Press. Pp. xvi+281. $47.00 (cloth); $19.00 (paper).

Frank Dobbin
*Princeton University*

The conventional wisdom about political conflict is at issue here. It presumes a world comprised of interchangeable actors, each of whom would behave as their adversaries behave in the place of those adversaries. It presumes a world of rational actors who pursue transparent self-interests, calculating their expenditures of political capital using a universal arithmetic of cost and benefit. Long before rational choice theorists distilled this view, it had gained wide acceptance among students of politics.

This view presumes that actors readily identify their objective interests and readily identify the best way to achieve them. It presumes that actors hold culture and identity at bay when making political decisions.

Wendy Nelson Espeland takes none of this for granted. At the foundation of *The Struggle for Water,* we do not find the assumption that political behavior is transparently and unproblematically rational. Instead, we find the assumption that rational behavior is transparently and very problematically tied up with culture and identity.

That this book is readable, engaging, and entertaining is surely to its credit. What could make better political drama than a group of federal bureaucrats fighting for a dam that will turn the better part of a huge Indian reservation into a lake? But what makes this book important is that it aims its arrows at the very heart of the matter in contemporary politics, the universality of rationality. Espeland examines the institutional context and ideas behind the behavior of three key groups in a pivotal controversy over land and water in the American West. What she finds is not three groups of calculating actors predictably pursuing their self interests, but three groups of actors with entirely different visions of what their self-interests are, of how to behave rationally, and of what is fair and right.

A writer might expect readers to accept the assertion that the Yavapai Indians whose land is at stake will not hold the same view as the federal engineers who drew up blueprints for a dam. Espeland's first brilliant stroke is to begin not by comparing the Yavapai with federal engineers but by comparing two groups of federal agents: the first, a group of engineers that has championed the dam on its technical merits for 40 years, the second, a new guard of bureaucrats that has a political-accomodation model of achieving consensus. The first group believes that the project can and will succeed on its technical merits. The second group has been

schooled in a model of controversy management in which achieving consensus is a political rather than a technical matter. These two groups could not be farther apart in their visions of rationality, but for each, professional identity drives how they conceive of problems and solutions.

The Yavapai are not much interested in the technical merits of the project—in its substantial benefits and modest costs. They are not much interested in negotiating a price that will allow them to think well of the decision-making process. They consider the land to be a part of their heritage, a part that cannot be bought or sold. Whereas the identities of the two groups of state agents are tied up with distinct professional models of decision making, Yavapai identity is tied up with the land itself. What makes the Yavapai most interesting is that they are hardly an isolated tribe untouched by modernity but a group that has self-consciously reinvented its own heritage in the image of multiculturalism. They appear to see their inalienable connection to the land, at least in part, as a product of their own construction.

Observing these groups interact through Espeland's narrative is a bit like watching *Who's on First*. They constantly misconstrue one another's meaning. The New Guard bureaucrats presume that the Yavapai are bargaining for more money when they say that their land is not for sale. The Old Guard federal engineers cannot comprehend Yavapai indifference to the benefits of the proposed dam and insist on repeating the same points ad infinitum. And the Yavapai seem baffled that federal agents know so little about current constructions of Indian culture that they do not grasp that the land is not a commodity.

The implications of Espeland's rich story of political struggle are wide-ranging. By documenting the complex connections between identity and conceptions of rationality, Espeland challenges simplistic models that read group interests and strategies directly from material positions. Each of these three groups has a modern conception of interest and rationality—one linked to a technical view of the world, another linked to a political problem-solving view of the world, and the third linked to an express multicultural view of the world. Espeland's argument is not merely that interest can be read from status characteristics, such as ethnicity and profession, but that groups' notions of their interests and of what is rational develop as their identities develop.

*The Common Place of Law: Stories from Everyday Life.* By Patricia Ewick and Susan S. Silbey. Chicago: University of Chicago Press, 1998. Pp. xvii+318. $42.00 (cloth); $16.00 (paper).

Michael McCann
*University of Washington*

This volume is the latest, and in many ways the most ambitious and impressive, contribution to sociological studies of legal meaning and prac-

tice in everyday life. Ewick and Silbey ground their study in an understanding of law as a constitutive force "in," rather than apart from, society. The authors' effort to identify law as "commonplace" underlines that all citizens are legal agents who act upon legal understandings and expectations, even when no official agent is present (p. 20). This emphasis leads Ewick and Silbey to offer a critical conceptual distinction between "law" and what they call "legality." Whereas the former term is used by the authors to refer to the formal legal authority of state officials, the latter is a broader term used to refer to "the meanings, sources of authority, and cultural practices that are commonly recognized as legal, regardless of who employs them for what ends" (p. 22).

The authors contend that legality is manifest in forms of citizen "consciousness." Consciousness, in this sense, refers to modes of cultural practice, to practical participation in the structures of social life; consciousness thus is not separate from or an effect of social structure, but rather "is an integral part of it" (p. 224). As such, Ewick and Silbey adopt a dialectical view in which individual agents and social structures are interrelated and mutually defining. Cultural codes or "schemas" that infuse human consciousness less determine subjectivity than provide constraints and opportunities for thought and action. And it is the specific forms of legal consciousness that define the focus of the authors' study.

The empirical project itself is grounded in 141 (of an original 430) in-depth interviews conducted as part of a seven-year project initiated by the New Jersey Supreme Court Task Force on Minority Concerns. Attention to stories and narratives is an especially critical dimension of the study. Stories provide at once objects and methods of study; they offer both a "lens" through which to inspect law in everyday life and a "metaphor" to represent what is discovered (p. 29).

Based on these narrative-oriented interviews, Ewick and Silbey identify three general types of commonplace accounts or forms of legal consciousness among their subjects. The account of "Before the Law" (chap. 4) envisions the law as if it were an alien, confining force imposed on society from without. Its archetype is the reified image of Kafkaesque "bureaucracy," in which law is rational and fixed, "discontinuous (and) distinctive, yet authoritative and predictable" (p. 47). By contrast, the form of consciousness specified as "With the Law" (chap. 5) envisions law as familiar, accessible, and manipulable—that is, as commonplace. The relevant metaphor in this case is an organized game among interested players. Finally, consciousness that is "Against the Law" (chap. 6) imagines law as a largely arbitrary and capricious power. Is it not distant and predictable, nor is it accessible and manipulable. Rather, this account imagines law as like a net of power that one seeks to avoid or escape. Its archetype is "making do."

The authors further maintain that these three modes of consciousness are not entirely separate or even opposed, but instead are interrelated in practice. This is important for two important arguments advanced in the book. First, the contradictions and tensions sustained by the presence of

the first two forms of practical experience sustain the very durability and flexibility of legality as a dynamic feature of American life. Second, the very practices of resistance against the law, and the potential for liberation, derive from recognition of those irreconcilable connections between particular lives and general social organization, and between the actual and ideal in law (p. 244). Building on these arguments, the authors develop their basic insights regarding how narratives of legality can both sustain *and* challenge hegemony.

*The Commonplace of Law* is a bold work whose analytical frame and methodology builds on but goes well beyond other recent sociolegal scholarship. Ewick and Silbey's conceptual contributions—including their distinction between law and legality, their identification of three forms of legality, and their discussions of hegemony and resistance—are significant. In particular, the authors' intriguing argument regarding how the contradictory character of legal consciousness sustains its very power and durability in social life is intriguing and persuasive. Finally, the book is clearly organized, eloquently written, and full of interesting discussions that animate the larger analytical project.

However, this reviewer was unconvinced or frustrated by some features of the book. Methodologically, the volume's many narratives of ordinary citizens' "thinking and doing" did not always clearly illustrate the larger theoretical points about legality, and especially about resistance against the law, made by the authors. More substantively, the authors' insistence that legality is wholly separate and unmoored from law seems overstated. While there is much to be gained by viewing legality as relatively more polyvocal and broadly ranging than official law, the authors' own discussions repeatedly confirm that the two are loosely tethered and interactive. At the same time, the authors' conceptual decoupling arguably leads them to eschew questions about how legality might tend to vary in relation to different substantive domains or institutional arenas of official law as well as among differently situated (by race, class, gender, etc.) citizens. In sum, even readers who share the authors' commitment to the study of law *in* society may be disappointed by the book's relative inattention to critical components of the diverse social contexts in which legality is embedded and emergent.

But these latter concerns perhaps are best framed as challenges for those who undertake further research of this sort. Ewick and Silbey have written a very thoughtful, insightful, and important book that is sure to inspire such further inquiry.

*The Myth of Green Marketing: Tending Our Goats at the Edge of Apocalypse.* By Toby M. Smith. Toronto: University of Toronto Press, 1998. Pp. ix+187. $50.00 (cloth); $19.95 (paper).

John C. Leggett
*Rutgers University*

Robert Merton and Karl Mannheim would applaud this work as an exercise in the sociology of knowledge. It is an effort to show how current day progressives, and the idiocies of false consciousness, are rooted primarily in material circumstance. For those of us weaned on the existential ties between base and superstructure, this book is trailblazing. It takes up dimensions of working people's consciousness and gullibility.

For a good contemporary example, consider the many blue-collar and lesser white-collar Riverside, California, residents who are sickened by pollution but lack the power to rename haze for what it is—smog. Working people must live with omnipresent green haut-bourgeois advertising copy produced to obscure the brown pall by bagging it with imperial scent: "The Inland Empire."

The metaphor constitutes a positive transfer which helps to anesthetize some people to the hideous ecology. However, there have emerged protestors against this green marketing hucksterism. And Smith has a lot to say about dissenters and radicals, plus their clever tormentors—the bourgeois green-thumb environmentalists. The latter class fraction distracts and dissuades through manufacture and distribution of myths around production and distribution of toxins. In a brilliant dissection of the two-page IBM advertisement headed. "In a Changing World Some Things Deserve to Remain Just the Way They Are," Smith details how the IBM green productionist discourse has been smudged by IBM reluctance to clean up its soil and groundwater pollution that exposes many thousands of residents to carcinogens. Given the toxic chemical threat, many laborers are awake and fighting to save themselves and their families, as in central New Jersey, where workers both female and male have tackled Johns-Manville, Union Carbide, and American Cyanamid. Traditional discussions of unionized working people's consciousness have neglected the fight-back ecological attitude as part of an informed, militant, blue collar mentality.

Smith shows how those who oppose the corporate ideologists have accumulated a lot of information from day-to-day conflicts against the bosses. In fact, there may be several material and ideational dialectics at work within the same objective situation: (1) owners versus working class, (2) extractive working class versus progressive middle-class environmentalists, (3) environmentalists versus the slick green thumbers, (4) the clever framers of the summary counter-conceptualization against most of the rest of us. Thanks to Smith, we can detail more clearly the attempts of the corporate elite to construct ecological hegemony. Remember, these are only attempts, not complete successes. Green thumb bour-

geois thought gets questioned in very popular films like *A Civil Action*. The fight-back message appears acceptable to the many who resist the brainwash.

More people from every class appear to see how global capitalist relations have pressed us into multiple catastrophes associated with ozone, acid rain, pesticides in the ground water, dead whales bloated with toxins, and Canadian silver seals stripped alive. The multinational paper industry generates mercury poisonings 1,500 miles west of the maritimes. They drain into the flesh of Ontario trout and nearby minamata Sioux. Thick glacial slabs of Antarctica—pieces the size of Connecticut—break away from their continental mass to cool the commercial fires of Borneo and soon to flood Manhattan's Battery Park.

Smith demonstrates how writers for bourgeois green ecology have disguised a globally sordid reality by way of propagandizing a simple theme—we can have our cake and eat it too. We former activist academics can shirk our citizens' responsibilities because by the late 1990s many of us had come to believe that the upper class, its scientists, and governments have the ecology problem under control. And it is all very democratic. The modern corporate elites have drawn upon consumer preferences, and not the need for high rates of capital accumulation, to promote allegedly less environmentally damaging products. Never mind that the substitute, for example, spray-on fiberglass used for the library ceiling, may be as damaging as the removed asbestos. Forget the librarian's eye pupil recently ulcerated by free-falling glass.

The capitalist greens have so far won in part because the authentic green movement does not know how to translate the bookish, research-based messages of the academy into the parlance of working people. So often the green intellectuals react against ecological outrages through manufacturing arrogant green messages toned at high levels of abstraction. As cynics and professionals, we allege to have a monopoly of unintelligible truth. Still we insist: our opaque arguments are the only ones worthy of an audience. And it is incumbent for people who find it difficult to read/accept our message to get in line. However, many of the nonbookish others do not want to line up. For often that means not just being run over by snobs, but losing jobs, occupations, homes, and regions they love. To talk to these people so different from ourselves, we have to listen and attempt to construct sentences with clear popular meanings and messages that make sense for familial material interests. In other words, we must learn from Jesse Ventura. That is the heart of Smith's message. Find a place for this splendid work in your course syllabi.

*"This Rash Act": Suicide across the Life Cycle in the Victorian City.* By Victor Bailey. Stanford, Calif.: Stanford University Press, 1998. Pp. xvi+349. $55.00.

Kevin Breault
*Middle Tennessee State University*

Historians have recently turned their attention to English suicide. The previous contributions to this effort are Olive Anderson's *Suicide in Victorian and Edwardian England* (Oxford, 1987) and Michael MacDonald and Terence Murphy's *Sleepless Souls: Suicide in Early Modern England* (Oxford, 1990). The present book is more narrowly focused on one city, Kingston upon Hull, 1837–99.

The authors of these books are committed to explaining suicide, and toward that end they have relied on various materials beyond history proper. Among Bailey's borrowings are quantitative description and Durkheimian theory. Note, too, that research methodology is an important part of the ongoing dialogue, as Bailey criticizes Anderson for her use of a small sample of inquests, and MacDonald and Murphy for a study that is "impressionistic and speculative, insufficiently grounded." What Bailey is saying is that the previous work is flawed for reasons of science. He is correct but does not fully appreciate how high now the standard is that he has invoked.

Bailey's book is a study of coroners' case papers for 729 inquests likely involving suicide, 604 in which a jury returned a suicide verdict. The data set appears to be a complete population of cases, the only surviving one from Victorian days.

Bailey sees the major debate in the sociological literature on suicide as one between the Durkheimian tradition and ethnomethodology, which views suicide data as socially constructed and highly unreliable. Bailey allies himself with the former and is especially supportive of the approach taken by Halbwachs, who focuses on the importance of urbanity and individual motives in the etiology of suicide.

Some minor points of criticism are that Bailey's discussion of the Durkheimian tradition is quite limited. Very little of the voluminous empirical literature of the last 10 years is considered. Second, too much is made of the ethnomethodological position, as it is virtually absent from the recent literature on the empirical study of suicide. Finally, one of the things we are reasonably sure about is that Halbwachs's emphasis on the city was mistaken. Rigorous individual-level studies have shown across sex and race/ethnicity that suicide is not more prevalent in urban areas (e.g., K. Breault and A. Kposowa in *Durkheim's Suicide: A Century of Research and Theory,* edited by W. S. F. Pickering and G. Walford [London: Routledge, in press]), a view which is strongly supported by historical scholarship during the Victorian age. Halbwachs's interest in individual motives is in the direction Bailey wishes to go (beyond Durkheimian ag-

gregate methodology), but the attempt is unsuccessful if only because the methods employed are inappropriate.

Suicide researchers will find chapter 6 on the incidence of suicide to be the most interesting. Among other things, Bailey presents suicide tables by year, decade, month, sex, age, age groups, method of suicide, occupational groups, and social class. From our standpoint, the most remarkable fact is how low the male suicide rates in Victorian England were. Even if we accept all of Bailey's suicides as genuine, the rate for males is only 12.2 per 100,000 (today, the rate for white males in the United States is 19.1).

Bailey finds strong evidence for social isolation (which he incorrectly refers to as Halbwachs's notion and not Durkheim's—how does one miss Durkheim's central idea?). However, all but rabid Durkheimians will find little comfort in the support offered by Bailey. What historical argument cannot do, what Geertz's "thick description" or Giddens's structuration theory cannot do, both of which the author notes favorably, is empirical science, the careful disentangling of covariates in controlled studies.

What would be the relationship between social isolation and suicide if the many variables discussed by Bailey could be properly controlled? Bailey admits that mental illness and alcoholism are involved in many of the suicides—are they causes of social isolation or vice versa? And if the often hidden factors of mental illness and alcoholism were controlled, would anything resembling Durkheim's theory be left? Bailey might be right, but unless this is a religion, I need more than anecdotal methods of persuasion, the psychiatric methodology of years ago.

What is required is an even thicker description where each case is put under the proverbial microscope, where a full set of social, economic, medical, and psychiatric considerations are carefully and quantitatively weighed and adjudicated. Historical study cannot help, nor can the traditional sociological approach in which suicide rates are analyzed. We must venture into the methodologies of social epidemiology, medicine, and research psychiatry if we are to adequately explain suicide and support Durkheim today. And I do believe support will be found.

For me, Bailey's book has come too late. As a beginning graduate student, I would have held it up as yet one more reason to resurrect Durkheim's theory over the prevailing resistance. With that battle behind us, Bailey's effort is only a reminder of the work that has yet to be done and the new resistance many sociologists have to doing it.

*Welcome to Middle Age! (And Other Cultural Fictions).* Edited by Richard A. Shweder. Chicago: University of Chicago Press, 1998. Pp. xvii+302. $41.00 (cloth); $16.00 (paper).

John Modell
*Brown University*

"Not everything real is something from no point of view" (p. xi) argues Richard Shweder in his introduction to this fine set of interdisciplinary essays about one such situatedly "real" "something": middle age. The book draws together specially prepared essays on mainstream and non-mainstream "middle age" (or its absence, or its closest approximation) in five radically different societies (including the contemporary United States), with an eye to examining "the emergence and diffusion of a cosmopolitan midlife discourse" (p. vii). That, at any rate, is the editor's thesis.

The essays' diversity—and the volume's internal controversy—is greatly heightened by a cultural-historical account by Margaret Morganroth Gullette of "middle-ageism" in the United States: the insistence that one *ought to be* anxious—perhaps preemptive—about incipient aging. (The text slightly preceded Viagra's entry upon the scene, but the drug and its uptake fit her argument perfectly.) "Belief in midlife decline is a pandemic peculiar to our era, a set of effects with a discernible twentieth-century history and extremely vigorous and insidious maintenance strategies" (p. 17). The "pandemic" as described is composed of "effects," but with no causes in sight; but these effects are said to have "maintenance strategies" but no one to formulate them is mentioned. Gullette rages against the diffidence of scholars, perhaps her fellow contributors, who launch "counterdiscourses . . . from . . . certain restricted 'fields' of academe (life course development, sociology, gerontology, midlife studies)" (p. 7). We are enjoined instead to construct better midlife narratives, for ourselves and as scholars, that "locate the decline elements in one's middle life more accurately" (p. 31), proximately in the producers of conventional decline narratives, more remotely in those who so arrange the labor market that early superannuation is in employers' interests.

Perhaps in reaction, Thomas S. Wiesner and Lucinda P. Bernheimer offer a spirited assertion of the meaningfulness of the notion of midlife in the contemporary United States, somewhat distantly justified by their fascinating empirical typology of life course deflections and reflections of a sample of American "countercultural" parents, studied longitudinally now for a quarter of a century. They also offer an empirically grounded assertion that seems to draw together the observations and arguments of several of the essays in the volume: "Midlife . . . is a joint, yoked transition of adolescent children and parents. Parents experience, define, and respond to midlife in part because of the teenage transition of their children" (p. 215).

Superb essays on Japan, India, and Samoa present sharp contrasts to

the American materials. Each rests on conventional ethnographic evidence, amplified by a bit of survey data, and constructs a clear and evocative sense of how being around 45 years of age fits into the scheme of things: each society has a dominant cultural account of the life course, but none has "middle age" in our sense.

Margaret Lock's comparison of female maturation in Japan and North America is especially powerful for its close attention to the body and to the systems of medical thought that construct what our exercised bodily capacities and our aches and pains tell us about our life stage and what our normative life stage tells us about what our bodies can and should do, ought and ought not feel. Lock's account, too, tells who it is that proposes and disseminates "cultural fictions" about middle age, in describing the operatives in the Japanese health care system that have latterly imported menopause or something like it.

Usha Menon and Richard Shweder write of women, too: Oriya Hindus. In no other essay in the book is the life course so fully or eloquently shown to be embedded in the reproduction of the family but in Orissa, where ideal households are patrilocal and generationally extended. Parturition and its cessation are not central to the adult female life course but are the passage into and out of the household of residents' wives. Menon and Shweder make powerful exposition of the trials (even somatic) of the new junior-generation wife and the relative comfort and honor that a woman of around 45 would likely experience.

Bradd Shore's cultural account of Samoa differs somewhat in its author's greater reliance on semantic analysis and preference for terse characterization rather than extensive immersion in Samoan culture. In this vein, Shore is able to explain the relative valuation that Samoans award to different parts of the life course in terms of the ascription to them of different degrees of activity on the one hand and power on the other. Mature adulthood in Samoa represents a decrease in activity and a sharp increase in power, and is very *different* from what precedes it, but it is no crisis of immediate or incipient decline.

Probably few sociologists today will be startled to learn that "middle age" differs across time and place, but most will value the opportunity to position themselves on what kind of a "cultural fiction" it may be. The evidence offered for the spread of a "cosmopolitan midlife discourse" is weaker, and the methodological challenge of establishing such a trend is hardly even broached. The importance and face plausibility of the thesis, however, no less than the excellence of the individual essays, makes this a valuable volume.

*The Gift of Generations: Japanese and American Perspectives on Aging and the Social Contract.* By Akiko Hashimoto. New York: Cambridge University Press, 1996. Pp. xv+226. $17.95 (paper).

Emily M. Agree
*Johns Hopkins University*

It might seem at first as though Akiko Hashimoto is indulging in stereotypes when she begins her book *The Gift of Generations* with the words of a lonely 74-year-old American woman who lives in public housing at least in part because her own children have died and her grandchildren ignore her. Yet, as interviews with the residents of West Haven, Connecticut, and Odawara, Japan, unfold across the pages of this engaging book, it becomes clear that her view is far more nuanced. A portrait of old age and family support emerges that reveals that there are many forms of intergenerational support in both the United States and Japan that involve both a reliance on cultural rules and also a process of negotiation among generations.

Hashimoto draws upon anthropological and sociological theories about the role of gift exchanges, both ceremonial and supportive, to develop her own perspective on the nature of the social contract between generations, using evidence from ethnographic fieldwork, surveys, and archival research to support her ideas. By raising the question of obligation to the symbolic level, she is able to apply her ideas about intergenerational support not only to generations within the family but also to cohorts within society as embodied by the social welfare system in each country.

According to Hashimoto, the central distinction between the two countries is in the way in which obligations for support are assigned. In Japan, obligations are culturally assigned according to family position, while in the United States they are more likely to be individually assigned according to an assessment of past giving or the promise of a future donation. Support of parents is a *duty* for Japanese children but a *voluntary contribution* for American ones.

Differences in the perception of support are derived from an understanding of aging itself and a differential definition of "deservedness." In Japan, old age is viewed as a distinct role defined by vulnerability. Intergenerational helping exists to prepare for the inevitable declines that accompany aging. In this context, the aged are "deserving" of the help that will guarantee security in old age. In the United States, on the other hand, old age is increasingly seen as an extension of independent adult life. The emphasis on self-sufficiency as a definition of adulthood means that help is negotiated on the basis of specific needs rather than as a blanket response to aging. It also means that all adult generations are "deserving" of independence, and there exists an obligation to self-sufficiency in old age in order to guarantee the freedom and opportunities of younger generations.

Yet, despite the definition of the Japanese contract as normatively

based, Hashimoto makes it clear through both examples and interpretation that the Odawaran parents *earn* the support of their children individually through "past sacrifices and hardships" (p. 177) and by making contributions to their children's household. The normative contract depends not only on moral obligations but also on direct reciprocity between parents and children. At the same time, the rules and assumptions that make an American intergenerational contract also are difficult to generalize. The United States is a multicultural society, and even among the Italian and Irish immigrants of West Haven, Connecticut, the assumptions about obligation and entitlement are quite varied.

The author does an excellent job of identifying the concepts and cultural assumptions that define the contract between generations, though the multidimensionality of family support in both cultures tends to make it difficult to consistently situate Japanese and American families in separate and unique systems of exchange relations. Instead, Hashimoto draws upon concepts from anthropological and social exchange theory to examine the *strategies* used to maintain exchange systems in each culture, including reciprocity, legitimization, empowerment, and equity. *The Gift of Generations* therefore provides an important synthesis of the diverse ideas that have been developed to explain giving behavior within families.

Hashimoto is a skilled interviewer with a unique ability to draw out much of the complexity in each older person's situation. However, her interviews are restricted to one generation's perspective. In only one case is a second person in the family interviewed, and that is a spouse. I found myself wondering many times how the picture of cultural rules and family obligations would change if multiple generations had been included in this study.

The interviews used in this book were conducted from 1982–84, which means that they predate many important social and political changes in both countries. In the United States, the last 15 years have seen a dramatic growth in home health care and changes in the rights of the elderly and disabled. In Japan, the Golden Plan, enacted in 1989, expanded the availability and acceptability of alternative forms of old-age support in Japan. Perhaps Hashimoto will return to these communities after 20 years and find out how the children of her original subjects are faring in this new environment.

*The Myth of Self-Esteem: Finding Happiness and Solving Problems in America.* By John P. Hewitt. New York: St. Martin's Press, 1998. Pp. viii+285. $29.95.

Colin Wayne Leach
*Swarthmore College*

John P. Hewitt's *The Myth of Self-Esteem* is part of a St. Martin's Press series on contemporary social issues. There is little doubt that Hewitt's topic is of general concern, as managers, teachers, parents, and politicians endorse self-esteem as an explanation of and answer to a wide variety of problems. Hewitt argues that this interest in self-esteem provides an opportunity to examine the particular "ethnopsychology" that produces it. He attributes U.S. investment in self-esteem to the deeply ingrained values of individualism, self-reliance, and self-improvement. The limited cross-cultural research on the topic does suggest that people in the United States are somewhat more concerned with individuality and self-worth than those in Asia, for example, but U.S. individualism is perhaps too easy an explanation. There are, in fact, many possible explanations, cultural and other, that range from norms of expression to politics to economic structure. Unfortunately, little effort has been made to examine more precisely the "cultural" explanation of the exalted place given self-esteem in U.S. society, and Hewitt's analysis does not help in this regard. The cultural explanation of self-esteem's popularity also fails to clarify why it is that self-esteem appears to have gained ascendance as a social explanation in the last decade or so. Is individualism on the rise, or is it conservatism, economic prosperity, or the growing divide between the chronically poor and everyone else? U.S. culture, writ large, is too general and too ubiquitous an explanation for the cache self-esteem seems to have in contemporary discourse.

Hewitt's theoretical approach is constructivist in orientation, with some use of symbolic interactionist and poststructuralist approaches to identity and individuality. He calls self-esteem a cultural myth because it is "a solution in search of problems" (p. 90), a general belief in self-esteem both frames problems and provides answers to them. For example, academic failure can be explained by students' low self-esteem, an explanation that leads directly to efforts to raise self-esteem. Use of self-esteem as explanation and intervention works to make the myth a reality, as people use the myth to make sense of their lives and to direct their actions. According to Hewitt, the myth is given "scientific" legitimacy by social scientists who "wholesale" it to the "conceptual entrepreneurs" writing popular books, appearing on talk shows, and consulting to businesses and schools. Why so many are convinced by the myth of self-esteem is left somewhat ambiguous, although superior marketing and cultural consistency are suggested as possible explanations. Hewitt is also uninterested in why social scientists, especially psychologists, persist in

research on self-esteem, despite its limited ability to explain psychological well-being or success in school or work.

After spending five chapters exploding the cultural myth of self-esteem, Hewitt, in a strange twist, uses the last chapter to claim its universal reality. He argues, with passing reference to physiology and human evolution, that "self-esteem . . . is both universal and culturally specific, hence it is both reality and myth. Its proponents think they have discovered something that they have in reality invented. But their invention of self-esteem and its incorporation as a key element of the American psyche is possible only because of universal features of human nature" (p. 123). These universal features of human nature include self-consciousness and a need to evaluate the self by relevant cultural criteria. The connections drawn between self-consciousness, self-evaluation, and self-esteem are given little support with argument, example, or evidence. In addition, little mention is made of the serious debates on these topics in philosophy, psychology, sociology, or anthropology.

The Myth of Self-Esteem is designed to fall somewhere between monograph and trade book, attempting to be both scholarly and accessible. It is written in a clear and relatively engaging style, with illustrative examples drawn from television talk shows, teachers, self-help books, and popular films, like the Wizard of Oz. Excepting the relatively detailed analysis of Dorothy's, the Tin Man's, and the Lion's search for self-esteem, however, Hewitt's examples have a generic quality about them. The absence of ethnographic or textual particulars make many of his examples appear rushed, with insufficient detail to persuade the critical or well-informed reader. Despite the author's admirable effort to appeal to both popular and scholarly audiences, the book is better suited for the former, in whom it may encourage a much needed skepticism regarding the importance of self-esteem and its ability to promote happiness and solve social problems. Given the low level of public discussion of issues for which the social sciences have relevance, Hewitt's The Myth of Self-Esteem may serve as an introduction to the constructivist view of social psychological ideas such as self-esteem.

The Fabric of Self: A Theory of Ethics and Emotions. By Diane Rothbard Margolis. New Haven, Conn.: Yale University Press, 1998. Pp. xi+207.

Rom Harré
Georgetown University

Foreigners who come to work in the United States, particularly those whose jobs are at the universities, at first find that there is much in common with the ways of life of their home places. But soon one profound difference emerges, particularly, I think, for the British. It has to do with the role of the institution in the life of the individual. With some exceptions, institutions are places in which people whose identities are forged

elsewhere pursue their individual careers. In other cultures, even those that are very closely related historically, it is the institution that is the source of identity, and careers are subordinated to its progress and honor. In this fascinating book, Diane Margolis brings to light some of underlying dynamics of modern ways of life through which we can gain an insight into the distinctive character of American mores.

The analysis and supporting arguments are built around something called "the self." I put the matter this way because the word "self" has acquired, again particularly in the North American context, a special range of meanings. (See, e.g., Charles Taylor's use of the word in *Sources of the Self* [Harvard University Press, 1989]). If one is to grasp the full depth and force of Diane Margolis's study, something needs to be said about her usage of the word. For much of what she says, something like "kind of person" would do as a paraphrase, but in some places the word seems to mean something more like "style of life." Of course this kind of person is inclined to live this style of life, by and large, yet there are important issues about how much the milieu demands of people of one sort—a form of life that is not what they would have led elsewhere. In the North American context, "self" sometimes means the attributes of a person, sometimes the kind of person someone is. And of course these are related but not identical. Bearing all this in mind, we find in this very well-written book a great deal of subtle analysis and considerable enlightenment into how life is currently being lived.

According to the author there are three main "selves": the "exchanger self," the type of person who flourishes in the market place; the "obligated self," the kind of person who willingly sees to the maintenance of the very substance of life itself; and the "cosmic self," the kind of person who seeks an identification with something transcendent to the everyday. Not only are these kinds of persons, but they are also systems of motivation, in the sense that C. Wright Mills gave to that concept, that is, not drives but explanatory schemes. One of the great virtues of the book, which contributes enormously to its readability, is the use of works of fiction to highlight the three-fold characterology that might otherwise be too bleakly abstract. The "exchanger self" is exemplified in the person of Soames Forsyte in the Galsworthy novels, the "obligated self" in the characters of Miss Eva and Mattie in Gloria Naylor's *The Women of Brewster Place,* and the "cosmic self" in the story of Siddhartha by Herman Hesse.

The key to the argument is in the historical balance between the necessities that link these "selves" into human life ways. The exchanger self, according to the author was more or less "invented" in the Renaissance (Hobbes gets some credit for it) and brought with it the conception of the social world as market place. However, human biology, bracketed by birth and death, is ubiquitous, *but has no place in the world as market.* So somehow the birthing and nurturance of new generations of exchangers must depend on the persistence of the obligated self as a kind of person and a life form that is realised in the role of woman as mothers.

The eventual dissolution of the market on the death of the marketeer makes a place for cosmic selves, life ways that draw their motivations from attempts at mergings, either with the Cosmos or Nature. Obligated selves and cosmic selves existed as person types before the Renaissance and have persisted into the world as market place just because human biology has as yet remained much as it was, though Diane Margolis points out that surrogate motherhood, cryogenics, and so on may yet swamp biology in a universal market.

Anthropology takes center stage when the plot turns toward the role of gifts in the maintenance of social orders. I was very pleased to see that the author had come across the marvelous book by Annette Weiner (*Women of Value; Men of Renown* [University of Texas Press, 1979]) that revealed the full extent of the way gifts created the integrated social worlds of men (order on a local scale) and women (order on a cosmic scale). This allows the author to enrich her catalog of kinds of persons to include the "reciprocating self," the "called self," and the "civic self." Each of these serves the central moral purpose of the book, to set aside the exchanger self as the necessary center of possible forms of life. The book is rounded off by the use of the currently popular constructionist view of emotions to demonstrate that the ways of life associated with each different kind of person (selves) modulate the meanings of emotions, for which we have seemingly universally applicable words, such as "love." This part of the book is somewhat thin, since there are all sorts of linguistic issues that are slid over.

As every sociological treatise should, this book takes a moral stance. The exchanger self is a defective type of person, relative to the possibilities for enrichment that are inherent in the human social organism. Anthropology is sketchily but tellingly introduced to show that the way the necessary role of obligated self is mapped onto persons, recently almost exclusively on to women, is ad hoc. There is no necessity about it.

The book is full of well-turned lines. For instance (pp. 47–48) "perennial complaints in the United States that 'it's not what you know but who you know' signal the continuing relevance of social ascent by biological descent." All in all, this is a work of the greatest interest, charmingly written, backed up by solid scholarship, even erudition, and to be warmly recommended not least because the author takes a stand on the moral qualities of different forms of life.

*Cultural Software: A Theory of Ideology.* J. M. Balkin. New Haven, Conn.: Yale University Press, 1998. Pp. xii+335. $35.00.

Susan Silbey
*Wellesley College*

After 250 years of writing about ideology, it is difficult to have something new to say that advances our understanding of this elusive concept, and

yet *Cultural Software: A Theory of Ideology* by J. M. Balkin manages to
do just that. Perhaps one of the most debated concepts in all of social
theory, there are dozens of definitions of the term "ideology," ranging
from relatively neutral uses describing sets of beliefs, ways of talking,
worldviews, and intellectual frameworks to critiques describing the false
consciousness and mystification of popular consciousness in the service
of power and injustice. Sociologists have sometimes described ideology
as the capacity to advance the political and economic interests of groups
or classes, or alternatively, the capacity to produce cohesion and resolve
social strain. Balkin attempts to move beyond these neutral, pejorative,
or functionalist conceptions by identifying underlying cognitive cultural
processes that produce ideology. Whatever the normative valence, the
use of the term ideology is usually meant to describe the ability of ideas
to influence social and material circumstances. Ideology is the effect of
what Balkin calls human cultural understanding, and in order to grasp
how such effects (ideology) may be produced, he offers a theory of how
cultures produce shared understandings as well as disagreements, how
they persist over time and yet evolve.

Balkin's major contribution lies in his specification of the mechanisms
of cultural cognition delineating the bridge between the individual and
the collectivity, between human mind and historical memory, between
micro- and macroprocesses. By invoking the metaphor of computer soft-
ware and its now commonly understood relationship to hardware, Balkin
suggests a model of the elements of culture in relationship to each other—
material technology, social institutions and know-how, or what he calls
the software. Cultural software refers to the tools of human understand-
ing by which we create other tools such as technology and human institu-
tions as well as more software; cultural software is a "toolmaking tool."
Cultural software is not a supraindividual entity or collective conscious-
ness, but rather "the collection of partly similar and partly different indi-
vidual copies" of "a conceptual apparatus within each individual that
prefigures cultural understanding but that can also change and evolve
over time" (p. 14). It is the operating system we each use, and sometimes
adapt, as we make our way in the world, leaving traces of our passage
that sometimes affect, sometimes infect, and sometimes fall dormant be-
fore future uses.

Deploying the metaphor of a tool, Balkin cautions, ought not suggest
some narrow instrumentalist conception. This is a work drawing much
of its sustenance from social constructivist traditions but one which
seeks to better specify how persons and institutions and history are con-
stituted and evolve over time. Relying on careful linguistic differentia-
tions to mark his agreements and disagreements with prominent social
theorists (e.g., Gadamer, Foucault, Horkheimer, Adorno, Lyotard, Bour-
dieu) and cognitive philosophers (e.g., Dennett, Putnam, Searle, and Gard-
ner), Balkin devotes the first part of the book to the elements of cul-
ture (e.g., meme, such as how to eat a meal in a restaurant, or belief in
a divinity), their accumulation through bricolage and their change over

time through processes analogous to but not quite exactly like biological evolution.

The second part of the book addresses the subject of the subtitle. If cultural software is the operating system of human sociality, ideology refers to the effects of cultural software that "help sustain unjust social conditions, unjust social relations, or the unjust use of power" (p. 104). Thus, after claiming to move beyond traditional approaches to ideology, and after providing an imaginative and erudite account of culture and history through the metaphor of cultural software, Balkin ends up using ideology, as many others before him have, to point to injustice. Analysis of ideology becomes analysis of injustice (see Silbey, "Ideology, Power and Justice" in *Justice and Power in Sociolegal Scholarship,* edited by Garth and Sarat [Northwestern University Press, 1998]). Despite the familiar notion, Balkin advances theories of ideology by differentiating ideological mechanisms from ideological effects. This distinction animates the entire book, for in this move Balkin hopes to distinguish description from normative judgment. Judging the effects of communication or cultural software is and should be, according to Balkin, analytically separable from describing its mechanisms.

There are echoes of epistemological positivism in Balkin's insistence on differentiating generic cultural software from specific ideological effects, although this is not a perspective he takes, nor about which he has much to say. Indeed, if I were to seek a fault with the book it would be this: because it fails to engage many of the debates among sociologists who have worked with these issues, the theoretical synthesis that is promised is disappointingly postponed. Those familiar with Sewell's "A Theory of Structure: Duality, Agency, and Transformation," (*AJS* 98 [1]: 1–29) will note that Balkin's use of meme seems to reproduce Sewell's notion of schema. Both Balkin and Sewell, like Giddens and Bourdieu, are attempting to explain how cultural know-how persists, is reproduced, yet changes and evolves. The language of agency and structure central to sociological theory is not, however, within Balkin's lexicon. This is not surprising, however. Balkin is a professor of Constitutional law and the First Amendment at Yale Law School worried about what seems to be cultural limitations on our ability to understand and do justice. Despite this limitation, which is better seen as a theoretical opportunity, *Cultural Software* is a remarkable work that will be usefully read by a broad audience.

*Mediating Social Science.* By Natalie Fenton, Alan Bryman, and David Deacon, with Peter Birmingham. Thousand Oaks, Calif.: Sage, 1998. Pp. xi+182. $74.50 (cloth); $24.95 (paper).

Wendy A. Wiedenhoft
*University of Maryland, College Park*

Recently the topic of knowledge production has captured the attention of social scientists. The growth of independent research institutes and policy-oriented think tanks worldwide have pushed many social scientists within the university setting, where knowledge production has historically been privileged, to contemplate their position of power as knowledge producers. Epistemological questions are being raised regarding who produces knowledge, where knowledge production takes place, how knowledge is disseminated, as well as what knowledge actually is (that is, the "ways of knowing"). The authors of *Mediating Social Science* frame the topic of knowledge production within media sociology to explore how social science is represented in the British mass media. They specifically examine media representations of knowledge produced by both universities and independent research institutes.

While research studies of social science in the media have been undertaken in the past, few, if any, offer a holistic approach to the subject. The authors of this study synthesize the two key realms of media sociology: active audience reception and journalistic agenda building. Incorporating the complex relationship between social scientists and journalists into media sociology, these authors are able to develop a comprehensive picture of the processes underlying the representation of social science research in the media. The actual mediation of social science is described not only as a struggle between the power and control of knowledge between journalists and social scientists, but also as a struggle with the way audiences interpret social science research.

Methodologically, a combination of qualitative and quantitative techniques were employed by these authors to encompass the views of the three primary groups—social scientists, journalists, and audiences—under investigation. A preliminary content analysis of national and local media news sources (newspapers, magazines, radio, and television) was undertaken. One major finding from this content analysis is that across all disciplines very little social science research is reported in the British mass media (p. 39). A finding that is significant to the sociology community is that journalists rarely explicitly refer to the discipline of sociology when reporting sociological research. Only 18% of social science reporting directly mentions the discipline of sociology. As the authors state, "sociology tends to be perceived as a low-status discipline in the social sciences and is less frequently referred to as a source or by way of conferring validity and credibility on a report" (p. 37). Psychology is the most frequently represented discipline in this study, accounting for 32% of all social science reporting.

Across all disciplines, social scientists assume the role of arbiters rather than advocates in the media. Few social scientists actively search out media sources to publicize their research; most respond to an agenda constructed by a journalist. The authors of this study state that "social scientists are, in the final instance, subordinated to an agenda that is not of their own making" (p. 103). As a result, journalists control what kinds of social science research is included (or, more importantly, excluded) in the media. Journalists have considerable power in the process of selecting newsworthy social science research to disseminate to the public. Epistemologically, journalists work within a positivistic framework; they value knowledge that is neutral and objective and facts that can be empirically verified. Other journalistic criteria for newsworthiness include topicality, generality, and a preference for concrete presentation of facts (usually through quantifiable measures). Social scientists that conform to the standards of newsworthiness set by journalists are more likely to appear in the media. Social scientists that frame their research within the discourse of the media are more likely to publicize the knowledge they produce than social scientists that produce abstract, theoretical knowledge.

The question of how social science in the media is interpreted and understood by the public audience plays an important role in the mediation of social science. Are audiences passive receivers of media discourse, or do they actively read their own meanings into a media text? The authors of this study subscribe to the later position, but through focus-group research they discover that audiences reading social science reporting tend not to subvert the original meanings of the text. The encoding power of the author, in this case a journalist, is not challenged as audiences decode media text in line with the author's intended meaning. While audiences may not challenge the power of the journalist, focus-group research did find that audiences are suspicious of the authority or expertise of social scientists portrayed in the media. Audiences responded skeptically toward the knowledge produced by social scientists, often viewing social scientists as "experts" who are out of touch with the "real world."

*Mediating Social Science* is a provocative study worthy of the attention of any social scientist working in a media saturated society. Competition for research funding and public recognition between universities and private research institutes, not to mention the power to produce knowledge itself, are just a few reasons why the way social science is represented in the media is relevant. As the government cuts back funding to institutes of higher education, social scientists must turn toward the private sector to secure resources for their research. Those social scientists able to gain representation in the media will be more recognizable and thus may be more successful in the struggle to produce knowledge. *Mediating Social Science* is a valuable tool toward understanding the intricate relationship between social science and the media.

*The Future of Anomie Theory.* Edited by Nikos Passas and Robert Agnew. Boston: Northeastern University Press, 1997. Pp. xii+240.

John H. Gagnon
*State University of New York, Stony Brook*

Durkheim's thinking about anomie, which played an important but somewhat different role in both *The Division of Labor* (cited once in this volume) and *Suicide* (cited routinely) was the subject (beginning in 1937) of a two-decades-long series of attempts by Robert Merton to find a practical utility for a classic social theory in the United States. In a move of great inventiveness, he selected one strand from Durkheim's ambiguous usage in *Suicide* and constructed a formidable theory of the relation between social structure and deviance. Rejecting all prior individualistic theories of separate pathologies, crimes, and delinquencies, he replaced them with a grand umbrella theory that was to link all forms and rates of deviance directly to social structure. The theory in its later incarnations focused less on variations in social structure and more on social psychological adaptations to discontinuities between goals and means as they were felt by persons differently located in the social structure, but this drift was characteristic of all Merton's work after his espousal of middle-range sociology.

The sheer ambition of this project attracted a vast body of theorizing and empirical work during the 1950s and 1960s. It is said that various versions of Merton's writings on anomie were the most cited works in sociology in this period. Commentary on his theory was at first admiring and later more critical. By the 1970s, Merton's encompassing vision of the relation between social structure and deviance had fallen out of fashion. A condition of empirical and substantive anarchy settled over the field of deviance (though not over the textbooks) with some forms of deviance losing their deviant status at least within sociology and with new forms of deviance being invented and incorporated into the sociological lexicon. A resort to a more comfortable world of separate deviancies with separate causes and separate experts has emerged: a world in which juvenile delinquency, child molesting, mental illness, organized crime, rape, and alcoholism have come back into their own as separately explainable entities.

In this edited volume, there are eight essays, two written by the editors, which search for a new utility for the Mertonian theory of anomie primarily in the substantive areas of juvenile delinquency and crime. Some essays attempt to defend and clarify various aspects of Merton's conception of anomie and make a case for his relevance to contemporary theoretical issues in criminology, particularly "strain theory" (e.g., the separate essays by Agnew, Passas, and Cohen, and the joint essay by Cullen and Wright). There are two empirical essays on delinquency that are remarkably different: the first is a sort of theoretical hodge-podge in which the theories of Bourdieu, Coleman, DiMaggio, and Merton are applied to the qualita-

tive study of the delinquent conduct of homeless ("street") youth (Hagan and McCarthy). The second is a dense and rather dissertation-like quantitative analysis of longitudinal survey data of parents and adolescents that examines Cloward and Ohlin's theory of differential opportunity (Menard). Still in the criminological vein, but working from a different aspect of Merton's protean theory, Rosenfeld and Messner start with the goal of unbridled achievement, which is the core of Merton's American dream (and Durkheim's commercial society), and discuss the possibilities of criminogenesis as moral restraints are eroded by a marketized society. Vaughn's essay is the only deviation from criminology and attempts to use anomie theory in explaining the organizational conflicts in goals (production versus safety) that resulted in the normal institutional practices that produced the Challenger space shuttle accident.

Do these separate essays, regardless of their individual excellences and weaknesses, bring the Lazarus of anomie theory back to life? I think not. The fit between the substance of these essays and the Mertonian perspective ranges from the uncomfortable to the tortured. Only the essay by Cohen in this volume reflects the copious ambitions of Merton's original formulation, and even his own early work on class and delinquency (which was concurrent with Merton's 1950s versions of anomie theory) was parallel to and not emergent from the Mertonian conception. Each of these essays reduces the substantive reach of Merton's ideas, and all do some interpretive damage to his original conceptions (as Merton did to Durkheim). It is perhaps a necessary fate of grand sociological theories to be selectively misread and misrepresented as they are untimely ripped from the historical circumstances of their birth to serve the purposes of a contemporary search for intellectual ancestry. If the authors in this volume share a common error, it is an assumption that sociological theories have a immediate utility outside of the contexts in which they are generated. This volume is absent the full and systematic interpretative reading of Merton's ideas that would be necessary to determine their possible relevance to the present circumstance.

*On Durkheim's Elementary Forms of Religious Life.* Edited by N. J. Allen, W. S. F. Pickering, and W. Watts Miller. New York: Routledge, 1998. Pp. xi+224. $85.00.

Edward A. Tiryakian
*Duke University*

In the past 10 years there has been an unparalleled flourish in Durkheimian studies in Europe and in North America. Perhaps unnoticed, much of the scholarly activity—papers presented at meetings, publications, archival research, and the like—about Durkheim, his colleagues, and relevant aspects of his milieu has become a collective enterprise. In fact, it might be said that what is emerging across national boundaries

is a second Durkheimian *équipe.* The contemporary Durkheimians are drawn by a common interest in explicating and translating texts, some previously unpublished; establishing networks and correspondence between the master, coworkers, and colleagues; and, most important in my judgment, in bringing out the relevance of the writings of the first Durkheimians to our own contemporary setting.

The new group endeavor has activists representing several generations and many nationalities. Philippe Besnard in Paris started in the 1970s *Etudes Durkheimiennes,* which brought to light renewed attention to Durkheim's ties with members of the *Année Sociologique.* In the 1980s, Robert Alun Jones at Illinois took over responsibility for the bilingual *Etudes* and eventually wired Durkheim in cyberspace. Finally, in the 1990s, the institutionalization process found a new and hopefully permanent foyer in Great Britain when in 1991 W. S. F. Pickering, with other dedicated associates like Ken Thomson, Mike Gane, W. Watts Miller, set up the British Centre for Durkheimian Studies at Oxford. Frequent conferences on selected themes have been held, with publications taking the form of both single-authored and multiauthored volumes, a high quality series of "Occasional Papers" (e.g., "Durkheim, Europe and Democracy," "Individualism and Human Rights in the Durkheimian Tradition"), and lastly, starting in 1995, an annual volume, *Durkheimian Studies/Etudes Durkheimiennes.*

The work presently under review attests to a primary focus being contributions of Durkheim and company to the sociology of knowledge and to the sociology of culture, particularly in the broad context of religion as an institution and as a symbolic system that can mobilize collective action. The seminal work is of course *Les formes elémentaires de la vie religieuse* (1912), which has in recent years stimulated innovative scholarship, for example, Jeffrey Alexander's *Durkheimian Sociology: Cultural Studies* (1988) and Donald Nielsen's *Three Faces of God* (1999), as well as benefiting from a new translation by Karen Fields (*The Elementary Forms of Religious Life* [1995], hereafter *Forms*).

Stemming from a 1995 conference, *On Durkheim's Elementary Forms of Religious Life* has 15 wide-ranging essays, a splendid introduction by Pickering, and a multilingual bibliography. As an opener, Howard Murphy's "Spencer and Gillen in Durkheim" situates Durkheim's selective utilization of the rich pioneering ethnographies of Central Australian Aborigines at a time when anthropology was undergoing a paradigm shift from a declining evolutionism to a rising functionalism. This essay provides an apt case study of how presuppositions can lead different secondary analyses to different interpretations of the same data set.

"Effervescence"—the climax of *Forms* (bk. 2, chap. 7, sec. 3)—has in our age of cultural movements (those of 1968 and 1989 in particular) become one of Durkheim's major heuristic notions. Three essays deal with it from different perspectives. William Ramp uses effervescence as an ingress to the wider body of writings of Durkheim and Mauss that relate to differentiation and dedifferentiation of self and totality. Drawing from

the interpretation of Georges Bataille, Camp sees in Durkheim's description of collective effervescence a dialectical rather than causal or functional account of social life as "a headless process in perpetual oscillation between energizing violence and the recovery of order" (p. 147). N. J. Allen uses the notion of "effervescent assembly" to speculate further on the origin of social structures in terms of his "tetradic" theory of four-section systems. Dénes Némedi's "Change, Innovation, Creation" repays close reading as a methodological critique of a circularity in Durkheim's argument resulting from "a faulty genetic argument where creative functions were assimilated with integrative ones" (p. 169). Némédi's reflections on Durkheim's "dubious theory of creativity" and his failure to "provide a satisfactory theory of institutional innovation" (p. 173) are a stimulant for post-Durkheimian theorizing.

Another cluster pertains to the intellectual context of *Forms*. Robert Alun Jones ("Religion and Science") brings to light a confrontation in 1908 between Durkheim and his former history of philosophy teacher at the École Normale, Emile Boutroux. Though voiced intellectually, their difference expressed a perennial tension in the sociological study of religion, at least insofar as positivist sociologists seek to objectify religious beliefs (p. 51). Sue Stedman Jones extends the notion of "belief" raised by Robert Alun Jones to the discussion of "beliefs" in *Forms* as representation, symbolization, and idealization of psychic energy central to social life (p. 58); she goes on to indicate how Durkheim's discussion owes more to his teacher Renouvier than to Kant. Ivan Strenski (author of *Durkheim and the Jews of France* [1997]) provocatively takes issue with Robert Alun Jones's close linkage of Durkheim's *Forms* with the interpretation of Robertson Smith regarding the nature of sacrifice and its relation to society and the sacred. Strenski sees a greater reliance of Durkheim on the studies of his students Mauss and Hubert, in turn influenced by Mauss's "guru," the Indologist Sylvain Lévi.

Leaving aside other essays that touch more on philosophical and epistemological issues of limited sociological interest, Kenneth Thompson's "Durkheim and Sacred Identity" has more relevance in outlining a Durkheimian general theory of ideology as a theoretical legacy of *Forms*, weaving Durkheim's discussion of clan, totemism, and personality to poststructuralist and postmodernist concerns of fragmentation. Thompson further provides an innovative critique of linear secularization theories, for example in relating Durkheim's discussion of tattooing to contemporary ethnographies of tattooees or "body modifiers" who reconstitute a symbolic community that sustains a sense of total identity (p. 101).

While the volume as a whole is too advanced in contents and in price to be assigned as required coursework reading, it does present a useful introduction to the "new Durkheim studies" that have added much wine to classic casks.

*The Politics of Religious Apostasy: The Role of Apostates in the Transformation of Religious Movements.* Edited by David G. Bromley. Westport, Conn.: Praeger, 1998. Pp. viii+244. $62.50.

William Sims Bainbridge
*National Science Foundation*

This is a remarkably unified collection of high-quality essays by many leading sociologists of new religious movements (NRMs), looking at the role of apostates in arousing hostility toward "subversive" religions. Four authors cite the example of the Family ("Children of God"), the first group attacked by the anticult movement, which James Richardson calls "possibly the most controversial of the early new religions" (p. 182). The Family is a millenarian, communal amalgam of Protestant evangelicalism and the California counterculture, which currently has about 20,000 members, half living in about 750 communes in more than 50 nations. The Family experimented for a brief period with a ministry that offered physical as well as spiritual love to the emotionally needy people it was trying to evangelize. Among the thousands of people who were involved over the years, a very few have been recruited by the anticult movement to play the role of apostates, encouraging governments to attack the group. In nations as diverse as France and Argentina, mass raids on Family homes seized hundreds of children and jailed adult members on charges of child abuse. The children and their parents suffered greatly in these episodes, but in all cases they were eventually released when the legal cases against them evaporated. This example illustrates the potentially grave consequences of apostasy for NRMs and the fact that apostate testimony often proves to be false.

David Bromley defines apostates not merely as defectors but as former members with conflicted departures who affiliate with opponents of the NRM. Low-tension religions ("allegiant organizations") often lack organized opponents and possess the power to define defection in harmless terms. In contrast, high-tension religions ("subversive organizations") are often victims of a dispute broadening process, in which the interests of opponents and the groups' inability to control the terms of the debate encourage defectors to become hostile apostates and endow their stories with great rhetorical power. Thus, the fact that NRMs have angry apostates is not proof of their evilness but rather the result of a process of symbolic interaction rooted in the fact NRMs are in tension with the sociocultural environment.

Four authors extend Bromley's model. Emphasizing the apostate's struggle to redefine personal identity, Armand Mauss argues that leave takers of all types may emerge from all types of groups, but analysis of Mormon history supports Bromley's prediction that apostasy would decline markedly as an NRM reduces its tension with the sociocultural environment. Eileen Barker notes that high-tension groups are seldom entirely cut off from the surrounding society but possess both peripheral

members (recognized as such by the movement) and marginal members (who are ambivalent, including individuals in the process of quietly defecting). Noting that most people who leave an NRM remain sympathetic to it, Stuart Wright suggests four factors that might turn a small minority of defectors into apostates: possessing an intensely religious primary role (master status) that cannot easily be abandoned, being in an NRM despised by society so that becoming respectable requires renouncing it, leaving through coercive deprogramming or exit counseling that inculcates hostility toward the NRM, and joining a social network of opponents to the NRM. Daniel Carson Johnson cites historical examples to show that many apostates' stories are factually incorrect.

Five other authors consider the organizational context. John Hall and Philip Schuyer examine the People's Temple, Branch Davidians, and Solar Temple, concluding that NRMs are unlikely to engage in violent confrontations with outsiders unless apostates organize threatening attacks on them. James Richardson analyzes apostate-initiated legal cases that had very serious consequences for NRMs, including draining their material resources and causing them to change fundamental policies, even in cases when the courts ultimately found in their favor. Susan Palmer develops a three-phase model of organizational response to apostates' testimonies based on the history of the Northeast Kingdom Community Church: first conflict arises between the group and individuals who seek external allies; then conflict escalates toward open confrontation, often fueled by highly inaccurate information the group's opponents spread; finally, the conflict decentralizes and deescalates as the group satisfies external authorities' demands and disinterested outsiders play moderating roles. Anson Shupe traces three waves in the historical evolution of apostate roles, arguing that the anticult movement is no longer as dependent upon them to initiate action as it once was.

In the concluding chapter on the Rajneesh Movement, Lewis Carter tackles the difficult question of what credence to put in the accusations of apostates, analyzing the divergent viewpoints of believers, apostates, opponents, and ethnographers. Through triangulation from multiple sources, taking account of probable biases, the social scientist can to some extent rise above the conflict. Courts and the general public should be skeptical of apostates' angry accusations, following the principle that the accused are innocent unless proven guilty.

*Gatherings in Diaspora: Religious Communities and the New Immigration.* Edited by R. Stephen Warner and Judith G. Wittner. Philadelphia: Temple University Press, 1998. Pp. vi+409. $59.95 (cloth); $24.95 (paper).

Matthew P. Lawson
*Brandeis University*

Religion's most important role in Stephen Warner's "New Paradigm for the Sociological Study of Religion" (*American Journal of Sociology* 98: 1044–93) is as a medium for expressing identity. Past waves of immigrants used religious communities to adapt and maintain their identities in the United States, and the situation is no different for those who have arrived since the 1965 Immigration Act. Yet "the religious presence of these groups, some 15 percent of the American population, has been ignored by scholars" (p. 6). This volume delivers 10 ethnographic studies of religious communities formed in America by new immigrant groups.

These immigrant groups left behind societies that maintained identity-producing differences and entered one where the stereotypes they are tossed into may not be to their liking. Some groups studied here use religious practices to resist these new distinctions. Catholic Mayan immigrants, for example, fled oppression by Hispanic people in Guatemala but must worship by their sides in American Catholic churches (Wellmeier). In consequence, they have formed intense parachurch groups that sustain their old ethnic bonds. Haitians in New York mark their differences from American blacks by publicly celebrating their Vodou-Catholic devotions (McAlister), while many Jamaicans do it in Rastafarian mode (Hepner). Affluent Iranian Jews give primacy to a Jewish identity over an "Arab" one (Feher). On the other hand, to form a numerically significant identity group at all, South Asian Hindus (Kurien) downplay caste divisions by raising caste- and location-neutral gods to greater prominence.

Religiously mobilized identity negotiations occur not only between immigrant groups and the host society but also within the immigrant groups themselves. Generation and gender roles are the chief constructions in question here, as six of the 10 essays demonstrate. As the second generation becomes Americanized and raises its status expectations, generational strains threaten congregational unity among Korean (Chai), Chinese (Yang), and Chicano-Mexican (Leon) Evangelical churches. The U.S. gender system has forced change in Rastafarian doctrine (Hepner) and Muslim worship practices (Abusharaf). And generational and gender struggles are compounded for Orthodox Christians from India's Kerala province (George). Men in these patriarchal immigrant families have lost economic status to wives recruited from India as nurses, but they find substitute status in church positions. Second-generation daughters—the essay's author among them—upset this balance by struggling for equality in church roles. Fengang Yang brings into sharpest relief the multiple

simmering tensions that threaten to divide the ethnic, linguistic, and generational factions in a Chinese congregation.

The race, class, and gender environment in America affects every group, but other changes arise from the institutional context of American religion. In some cases (Rastafarians, Hindus, Yemeni Muslims), a congregational structure and professional clergy are adopted where none existed in the home country. In other cases (Korean, Chinese, Chicano-Mexican), immigrants adopt American Evangelical practices more often than they would at home. In the "new paradigm," this is not surprising: in the religious marketplace, entrepreneurs who adapt to the needs of constituents are more likely to win adherents, and in America the entrepreneurs are likely to be Evangelical.

Though the essays in this volume were commissioned by Warner to advance his "new paradigm," he gave the authors the freedom to address issues that arose from their research and in the literatures of their respective immigrant groups. This makes specific essays attractive to particular audiences, but it also promotes variability in the kinds of information each author reports. What is most lacking is attention to the wider context in which each community finds itself. Some essays benefit from additional research in their group's home countries (Wellmeier, McAlister, Hepner). But most pay insufficient attention to wider institutional contexts even in the United States (Kurien, Wellmeier, and Hepner are partial exceptions). If the competitive context is the cornerstone of the "new paradigm" (Warner 1993), then the reader would like to know about competing religious organizations that may serve these groups, as well as secular ethnic associations. Warner's introduction to the volume provides some of this contextual information and compares the religious activities of past waves of immigrants to this one. He also provides an overview of the research project and its aims. Judith Wittner's conclusion draws out the sociological significance of themes that arise in individual essays and in the collection as a whole.

This volume provides a set of fine-grained ethnographic studies of immigrant identity strategies mobilized in a religious medium. Those interested in ethnic identity, and identity construction more generally, will find the collection valuable. Sociologists of religion will be interested in the novelty of some of the groups described, as well as the novel uses of religious traditions they think they know well.

*Social Movements and Their Supporters: The Green Shirts in England.*
By Mark Drakeford. New York: St. Martin's Press, 1997. Pp. ix+236.
$65.00.

Ira Silver
*Northwestern University*

There has been considerable inquiry in recent years about what is the
most suitable framework for studying social movements. In probably no
subfield of sociology have debates about "micro v. macro" or "culture
v. structure" been so pronounced. Within this context, Mark Drakeford
examines an interesting case of a social movement that emerged in Brit-
ain after World War I. The Green Shirts evolved from a most unusual
source—youth scouting—and sought to link the outdoor way of life to a
broader agenda for social change. Led by John Hargrave, the movement
originally called itself the Kibbo Kift, changed its name in the early 1930s,
and gradually declined after World War II. Influenced by network ties
to organized labor, the Green Shirts' goal was to institute Social Credit,
a program for economic reform during the Depression years.

Following the work of Sidney Tarrow and Charles Tilly, Drakeford
seeks to capture the life cycle of the Green Shirts movement—its endur-
ing capacity to define problems, propose solutions, and carry out agendas.
He views a movement's work in fostering participation and loyalty as
dynamic in that both structural and historical forces condition this work.
Therefore, in doing a microlevel of the Green Shirts, Drakeford acknowl-
edges that this level of analysis alone cannot produce "an understanding
of how a specific movement emerges, or develops along a particular path"
(p. 3).

The book offers readers an excellent overview of social movement the-
ory through the lens of a thoroughly researched historical case. Drakeford
is sensitive to limitations in his data—particularly the fact that his inter-
views during the early 1990s of surviving members of the movement con-
tained recollections that were over half a century old. Nevertheless, the
biographical accounts he recorded of movement participation are rich.
From these interviews and an extensive body of archival materials,
Drakeford has substantial data that he meticulously weaves into a discus-
sion of social movement theory.

However, some concerns arise about how well he has done his home-
work. While he refers in depth to social strain and resource mobilization
theories, only one reference is made to the burgeoning literature on fram-
ing. He writes, "The changing character of the Kibbo Kift suggests that,
for social movements, such framing itself is continuously problematic and
more at the mercy of their own history and external events than theorists
often suggest" (p. 108). This sentence articulates a claim for which the
book subsequently provides considerable evidence. This is the idea that
personal motivations are necessary but not sufficient bases for forging
commitments to a movement. While a probing idea, a discussion is never

given of how the work of Robert Benford and others who work in the framing tradition fail to attend to the way in which external factors influence the microlevel work of movements.

In a related way, Drakeford does not consider the importance that collective identity may have had in shaping movement commitment. The important literature on the subject is not cited, yet there is ample evidence to suggest the importance of collective identity in the early Kibbo Kift years. The group had not yet made a' substantial transition from woodcraft and camping to wider concerns about social and economic change. While participation marked a celebration of youth, peace, and a natural way of life, in what sense were participants also celebrating themselves? Even with the adoption of the Social Credit agenda upon the switch to the Green Shirts name, participants may still have blended instrumental and expressive concerns.

Finally, it seems that Drakeford could have done more in theorizing about some of the unique features of his case. What, for example, is revealed by the fact that the Green Shirts began as a youth and cultural organization and became a militant political force? How does this transition compare to the life cycle of other contemporary movements? What were the consequences of having equal membership of men and women? In sum, while the Green Shirts case offers a more nuanced understanding of the life cycle of a social movement, the case would have been even richer analytically had it been situated in a comparative framework.


*The Politics of Downtown Development: Dynamic Political Cultures in San Francisco and Washington, D.C.* By Stephen J. McGovern. Lexington: University of Kentucky Press, 1998. Pp. xiv+342. $44.95.

Dingxin Zhao
*University of Chicago*

This book explains different outcomes of downtown growth-control movements in San Francisco and Washington, D.C., the two cities that the author considers as strikingly similar in many aspects. The author showed that the growth-control movement in San Francisco was more successful because the activists over time had adopted an inclusive and grassroots based strategy of mobilization. The success of the movement brought positive changes to people's (including developers') understanding of the issues and to politics of downtown development. As a result, developers had to compete for the limited office development quota set by the law. The process raised the quality of downtown buildings and the profit of the investment, and therefore, led to a win-win outcome to both local communities and the developers. On the other hand, the growth-control movement in Washington, D.C., consisted of two competing groups (the planning advocates and community development advocates) with different views on the issue of downtown development and

the role of government in the development, and implemented a style of mobilization that was more or less exclusive and elitist. In the end, frictions between two organizations undermined the movement, and the activities of the growth-control movement organizers unintentionally favored a downtown development coalition.

The book is very clearly written and is full of interesting and compelling details. It is an important addition to the study of American urban politics and its mechanism of change. It also has contributed to cultural studies, which in the past, have derived individual behavior from cultural categories. After reading the book, I am once again convinced of the importance of ideas and strategies in social change.

The book is also guided by a theoretical paradigm centered on the concept of cultural hegemony. However, the effort was less successful than the empirical account. The theoretical framework (p. 34) appears to be more a narrative device than an analytical tool, without which the story has to be told differently. While the author criticizes resource mobilization theory (p. 44), the book can be read as arguing that a populist "mobilization structure" (or framing) is more successful than an elitist one for the downtown growth-control movement. I also feel uneasy about the author's use of the concept "cultural hegemony." Although there is no absolute dividing line, by cultural hegemony, what Gramsci had in mind was the domination of bourgeoisie political culture over drastically different political ideas. If a change of strategies (even if it brings changes to some social practices) under an interest group politics can be described as the rise of a counterhegemonic culture, we should expect that counterhegemonic cultures rise everywhere, and thus cultural hegemony is no longer a problem for Gramscians.

The author may have also underestimated the impact of other factors on the formation of strategies and outcomes of downtown growth-control movements in the two cities. My reading of the book shows that politics is much more racially divided in Washington, D.C., than in San Francisco. It was probably due to a strong racial division, among other factors, that most African-Americans supported Mayor Barry (a convicted cocaine addict) and his most inefficient government (that had one city employee for every 13 residents [p. 259]) and questioned Barry's legacy only after he left the office (p. 260). The fact that black and white neighborhood groups in Washington have cooperated on one occasion (p. 273) does not mean that it is easy for them to cooperate more extensively. Moreover, developers and city planners in Washington, D.C., never exhibited any ambition to the "Manhattanization" of the city as their counterpart in San Francisco did (p. 63). Between 1976 and 1983, San Francisco had developed a total of 21.7 million square feet of downtown office space (calculated from table 4.2, p. 80). In the same period, Washington only developed 14.8 million square feet of downtown office space (table 9.1, p. 199). If we assume people's resistance to downtown development is proportional to the scale of that development, a strong growth-control

movement in San Francisco may be attributable to the city's scale of development.

Having stated some reservations, I want to emphasize that this is a book out of a very logical mind. The author has convincingly demonstrated that ideas and strategies can not only have an impact on the outcome of a social movement, but can also change the view of the people and the further development of a social movement and society. I congratulate the author for presenting us with a compelling story with insightful details.

*On Becoming Homeless: The Shelterization Process for Homeless Families.* By Ione Y. DeOllos. Lanham, Md.: University Press of America, 1997. Pp. xxxv+235 pages. $56.00 (cloth); $34.50 (paper).

Joanna M. Badagliacco
*University of Kentucky*

The number of homeless families has increased in the United States, especially over the last two decades. Ione Y. DeOllos offers a preliminary account of a variety of families who are in housing distress and attempts to explain what happens to families to cause them to seek public shelter, how they live in shelters, how they view themselves, and how the availability of familial support affects this process. This general overview suggests a viable theoretical approach but at the same time limits her ability to offer a more in-depth discussion of particular issues.

For example, the subtitle of this book indicates an emphasis on homeless families, a much understudied population. Yet, I found that in many places DeOllos presents secondary data that refers only or mostly to homeless individuals or "singles," as they are apt to be known in the shelter community. This information may be useful as background but confuses the discussion of *families* who are homeless. For example, table 5-1 (p. 97) lists families as only 12% of the 8,617 residents during 1987–90 at one of the shelters she studied. Who were these families? And how is "family" defined? If families were the focus of this work, I would have welcomed additional information about what constitutes these families.

Another table, 5-3 (pp. 101–2), makes no mention of families at all but reports only on "individuals," including men, women, and children. This particular table is constructed from survey data collected by a former member of the shelter staff, who, according to the source notes, surveyed 132 individuals, or only 1.5% of the 8,805 persons reported to have resided there in 1990. Moreover, his methodology is noted as "not available" (p. 109 n. 2). This is, however, the same shelter mentioned previously as having 8,617 persons residing there from 1987–90. Now, just a few pages later, we are told that 8,805 persons lived there in 1990 alone. DeOllos is no doubt trying to give the most complete picture available, but she

should not include seemingly inaccurate and suspect data that only causes the understanding of the demographics of those homeless to be much more cloudy.

I think readers will find her chapter on the history of homelessness quite interesting reading that helps place contemporary research within a rich and longstanding context. Homelessness, however defined, is not new, of course.

Other highlights of this work are the interesting diagrams of what DeOllos calls the "household careers" of homeless families, that is, their paths from once being domiciled to eventually being sheltered. She provides illustrations of alternatives by describing the experiences of several families, although she does not tell us how or how many of the families choose one or another alternative "career"—information that could be valuable for policy and program development.

Five intriguing "parental shelterization process stages" are introduced in this work: (1) compliance, (2) anger, (3) hope, (4) frustration, and (5) hopelessness (p. xviii). These stages are drawn from the works of Edwin Sutherland and Harvey Locke (*Twenty Thousand Homeless Men* [Lippincott, 1936]), David Glaser (*The Effectiveness of a Prison and Parole System* [Bobbs-Merrill, 1964]), and Erving Goffman (*Asylums* [Doubleday, 1961]). Theoretically, these are exciting and important concepts that have the potential to aid in understanding the processes families undergo based on their interactions and relationships with staff, other families, and their own kin. Unfortunately, DeOllos only addresses these stages relatively briefly toward the end of the monograph. Her book would have been substantially strengthened if she had integrated and woven these ideas into all of her work and compared her observations with earlier models using different populations.

In chapter 8, DeOllos presents three hypothetical composite families, claiming that "stories of hypothetical families provide 'ideal' examples which demonstrate how families experience shelter life" (p. 143). I believe that she shortchanges readers by failing to take the opportunity to give these homeless families a voice and let them speak of their experiences in their own words. The poor, and especially the homeless, are rarely given voice by researchers; every effort should be made to allow families to speak for themselves. Their own words are empowering and meaningful. They help to clarify, are less likely to be misinterpreted, and are important. The families become real when they are allowed to speak for themselves.

DeOllos's work is a starting point for future work. She raises some possible approaches that could be considered by researchers when examining the pathways followed by families as they descend into homelessness. Yet, her own analysis of those pathways needs strengthening and expansion. For example, in looking at the paths taken by those families she met, she might examine strategies for moving on. Do families reverse the processes she introduces in order to become domiciled once again? In addition, how do issues of gender relations and race and eth-

nicity impact and influence choices? Several times in the text she mentions violence and abuse in families. How does family violence relate to interpersonal relationships, especially when examining kin support? And, what policy implications can she draw from her research?

One final technical note: readers are apt to find a lack of technical quality in the book. For example, the index is inadequate, editing seems hastly done, and some of the names used to describe individuals and places might be misinterpreted as disrespectful, and maybe even offensive.

*Black Corona: Race and the Politics of Place in an Urban Community.* By Steven Gregory. Princeton, N.J.: Princeton University Press, 1998. Pp. xii+282. $29.95.

*Hamilton Park: A Planned Black Community in Dallas.* By William H. Wilson. Baltimore: Johns Hopkins University Press, 1998. Pp. x+256. $39.95.

Mary Pattillo-McCoy
*Northwestern University*

The African-American communities studied in these two books—Corona in Queens, New York, and Hamilton Park in Dallas—occupy distant positions within the spectrum of black residential patterns. The black settlers in Corona had established a church as early as 1828, while the planned subdivision of Hamilton Park was only recently built (in 1954), and on previously unoccupied land. Home to Italians, Irish, Germans, and African-Americans, Corona was integrated from its beginnings through the 1950s, when it became nearly all black. Hamilton Park, on the other hand, was conceived of and implemented as a racially segregated *and separated* community, although subsequent migrations have placed it within a ring of white residential and commercial developments. Despite these distinctions of history and context, both Corona and Hamilton Park provide insight into realms of black life that have been eclipsed by a focus on "the black ghetto." The local actors, through their civic participation, counter the stereotypes of black communities as culprits of urban decline and liabilities in processes of urban regeneration.

In *Black Corona,* anthropologist Steven Gregory highlights the diversity of black life "by directing attention to the political struggles through which Corona's black residents have contested the practice of racial exclusion and, in the process, negotiated multiple and shifting meanings of race, class, and community" (p. 12). Gregory's project, then, is to understand black identity as it is shaped by place-based politics. In doing this, Gregory dispenses with class categories that equate socioeconomic characteristics with "mainstream" values or "respectable" behaviors, obscuring more than they clarify. Instead, Gregory explores black identity as it

is affected by the locations of particular groups of African-Americans in political struggles for community empowerment.

In the 20 high-rise apartment buildings of Lefrak City, for example, African-American parents organized against the derisive categories that white community leaders used to define them: fatherless families, irresponsible mothers, and (therefore) unruly children. Black residents protested not with words, however, but in their deeds. In organizing the Concerned Community Adults and its Youth Forum, which spearheaded neighborhood cleanup efforts and brought attention to the problems of police harassment of black youth, the black residents of Lefrak City repositioned themselves in the neighborhood discourse from being perpetrators to being political actors. Their participation challenged the discourse of black family pathology and black youth criminality.

The black homeowners of East Elmhurst (once a part of Corona but later symbolically separated by class distinctions) faced different problems and utilized different frames than the residents of Lefrak City. Threatened with plans for locating a foster care facility in their neighborhood, or the expansion of LaGuardia airport, leaders in East Elmhurst often invoked the importance of property values and "life investments," with a simultaneous understanding of how local issues "were intimately linked to racially structured patterns in the distribution of social resources and burdens" (pp. 155–56). The parallel in the activism of Lefrak City renters and East Elmhurst homeowners is that both insisted on self-definition rather than the impositions of outsiders. Yet the diversity of the black community is reflected in the variation in identity narratives: "law-abiding youth," for example, in Lefrak City versus "property-tax payer" in East Elmhurst.

In *Hamilton Park,* historian William H. Wilson is less concerned with definitional struggles based on race, class, or gender. Indeed, the subtitles of the two books are indications of their dissimilar analytic agendas. Gregory uses broad categories in his subtitle—"Race and the Politics of Place in an Urban Community"—whereas Wilson's subtitle—"A Planned Black Community in Dallas"—is straightforward and factual. Accordingly, *Hamilton Park* is more instructive than it is illuminating.

Wilson finds in the Hamilton Park subdivision a "community" of the sort believed to be destroyed by splintered allegiances and responsibilities. Following the Chicago school definition of "community," Wilson describes the place-based orientation of Hamilton Park families and a corresponding "self-identity grounded in its institutions and activities" (p. 9).

The history of Hamilton Park is informative in part because of the interracial leadership that created it. As in most cities, Dallas's black population lived in overcrowded and decaying parts of the city. Dallas's African-American leadership lobbied for solutions to the black housing problem, but the frequent bombings in neighborhoods experiencing racial turnover in the 1950s were the more immediate impetus for Dallas's white elite to get involved in providing housing for blacks in Dallas. White bankers, insurance executives, and oil tycoons worked with black

leaders to locate suitable land (i.e., away from white settlement), coordinate the project's financing, and construct and market the subdivision. Not long after its completion, however, the hand of white benefactors steadily receded and Hamilton Park began the business of institution building and organizing as a segregated community. The present-day situation of racial segregation, social class diversity, and organizational activism is where the stories of Hamilton Park and Corona converge.

Both authors chronicle residents' confrontations with encroaching industry and commerce, as well as contests over school desegregation. For example, the East Elmhurst–Corona Civic Association won concessions from hotel developers and lobbied for cleanup of Flushing Bay, while the Hamilton Park Civic League prohibited a lumber company from locating within its boundaries. Where the books diverge, however, is in their treatment of these accounts. Wilson uses examples of neighborhood activism to argue for the importance of place-based communities, although he offers little in the way of a sustained theoretical discussion and critique. Issues of race (and class and gender) are underanalyzed in *Hamilton Park*. For example, the interest that white businessmen showed for the black housing crisis goes unexplained beyond their liberal leanings (in the Southern context) and their concern for white neighborhoods in the path of expanding black areas. On the issue of schools, the predominantely white school district in which Hamilton Park was situated decided to close Hamilton Park's all-black junior high school rather than to integrate it, but Wilson's story provides only a glimpse of the heated racial politics that was likely to have accompanied such a decision.

Gregory provides these kinds of analyses in *Black Corona*. He investigates class within race, race within place, and place within politics. Gregory also pays close attention to relations of gender power. But if Wilson errs on the side of too little conflict in his narrative, perhaps Gregory places too much emphasis on conflict as the site of black identity formation and negotiation. For example, Gregory argues that Corona did not become a "black community" until it was "denied political power on the basis of race" (p. 65). Yet, in this reasoning, Gregory does not fully recognize his own data about the positive and proactive creation of black churches and social clubs by early black settlers in Corona. While Gregory makes important points about the structuring of black identities in reaction to racial, class, and gender hierarchies, he gives less analytic attention to the creative and generative facets of black life.

Both books provide a window into the diversity of the black urban experience. *Hamilton Park* is appealing for its depiction of a unique type of black residential development—southern, quasi-suburban, and intentionally segregated—while *Black Corona* succeeds in drawing parallels between the local case and a larger concern for understanding systems of stratification and processes of identity formation.

*Shifting the Color Line: Race and the American Welfare State.* By Robert C. Lieberman. Cambridge, Mass.: Harvard University Press, 1998. Pp. xiii+306. $45.00.

Yvonne Zylan
*University of Arizona*

In the history of American welfare state development, the role played by race and racial discrimination is an open secret: rarely discussed explicitly but widely understood to be profoundly influential. More often than not, this approach is replicated in scholarly research on the subject. Race is addressed only anecdotally or conflated with regional political interests (the South versus everyone else). Few scholars have attempted to systematically determine the extent to which race has shaped welfare state commitments, and fewer still have tried to use that analysis to theorize the racial dimensions of political life more broadly.

Robert C. Lieberman's *Shifting the Color Line: Race and the American Welfare State* is a welcome exception. Written for an audience of welfare state researchers, but equally accessible and valuable to a general sociology audience, Lieberman's book persuasively demonstrates that "race inhibited the development of a strong, unitary, centralized welfare state in the United States, and that the fragmented American welfare state helped to reshape the politics of race and the place of racial minorities in American life" (p. 6).

Lieberman mixes basic quantitative techniques with a carefully researched historical analysis to explain the development of three main components of the U.S. welfare state: old-age insurance (OAI), Aid to Dependent Children (ADC), and unemployment insurance (UI). He is not simply concerned with whether or not these programs served to discriminate against recipients and potential recipients based on their racial identity. Rather, he is interested in how the racially specific incorporation of Americans into social programs has shaped the possibilities for political, social, and economic inclusion. Given the centrality of welfare state politics in advanced industrial capitalist democracies, has racial stratification via systems of social provision reinscribed racial domination?

Lieberman's answer is yes. The political marginalization of African-Americans in the present day United States can be traced to the racial tensions embedded in welfare state politics. And yet, he notes, the welfare state has itself been variably constituted by race. Old-age insurance has developed into "the closest thing to a universal, color-blind social policy the United States has ever had," while Aid to Dependent Children has historically discriminated against African-Americans and has made them "the central focus of bitter political disputes" (p. 8–9). Unemployment Insurance occupies the middle ground—not the racially divisive ADC, nor the relatively benign OAI, but a program that has kept "African-Americans at the margins of the political economy" (p. 9).

Although Lieberman's account stresses institutional structure, it is not

an "origins" story. He notes that the original configurations of OAI and ADC, for example, would seem to have predicted a more racially charged and discriminatory future for OAI than ADC. OAI originally excluded the vast majority of black workers, while Aid to Dependent Children contained no explicit language of exclusion and in its early years opened the program to greater numbers of black families by requiring that it be available in every county of a state choosing to administer it. Yet over time, OAI served to include greater numbers of African-Americans, and debates over the program moved away from the politics of race. ADC, however, became the focus of racially inspired attacks and was debilitated by periodic waves of reform that sometimes explicitly singled out black women for punitive treatment.

What explains the different trajectories of these programs over time? Lieberman's answer is institutional structure and its dynamic interaction with group political processes. Specifically, he argues that OAI expanded its reach and its political support by virtue of the fact that it was a national program, while ADC suffered the consequences of parochialism: patronage-style implementation, political manipulation, and public mistrust. Unemployment insurance—a hybrid program—charted a middle course.

Lieberman's analysis is clear and convincing, and his use of quantitative techniques is deft. The chapter on ADC is easily the most compelling of the book, and the reader sometimes gets the sense that those on OAI and UI are there largely to buttress the arguments regarding this now defunct program. Of greater concern to welfare state scholars should be the nearly complete omission of any discussion of the gender dimensions of policymaking. Lieberman's claim that "race and class were mutually constitutive in the making and growth of the American welfare state" (p. 26) is likely to provoke many working in this area to ask "and what about gender?" as there is now a large body of work demonstrating its importance. To be fair, many gendered accounts of welfare state development do a poor job of incorporating race. Still, this is not a small point; one obvious answer to the question of why ADC has traveled the road that it has is that it provides benefits mainly to women, while OAI and UI are male-dominated.

That said, the book is well worth the attention of sociologists interested in race, politics, stratification, and public policymaking. It should also be read by anyone who wonders what the effects will be of current efforts to "devolve" programs of social provision to the state and local level. If the history of these three programs tells us anything, it is that the best hope for racially inclusive, generous, and politically popular policies is to be found at the national level. Failing a commitment to such policies, we doom ourselves not only to a future of welfare state squabbles but to the perpetuation of the color line in American politics and society.

*Charting Chicago School Reform: Democratic Localism as a Lever for Change.* By Anthony S. Bryk, Penny Bender Sebring, David Kerbow, Sharon Rollow, and John Q. Easton. Boulder, Colo.: Westview Press, 1998. Pp. xviii+383. $69.00.

Dennis Shirley
*Boston College*

Beginning in the late 1980s, the public schools in Chicago undertook one of the boldest efforts at school reform in the postwar United States. Long the site of many of the nation's most contentious school wars—best described in George Counts's *School and Society in Chicago* [Harcourt, Brace, 1928] and Julia Wrigley's *Class Politics and Public Schools: Chicago, 1900–1950* [Rutgers University Press, 1981]—Chicago has often been noted for its unusually intense politics of race and class in matters of education. Regretfully, we have lacked a thorough scholarly account of the de facto collapse of the schools in the postwar era and the loss of the public's faith in the ability of the city schools to educate their children. It was this critical state of affairs that provided the backdrop for a courageous reform, launched in 1988, to decentralize Chicago's public schools to expand genuine grassroots community control of the schools.

Exasperated by the condition of the Chicago schools and stung by former Secretary of Education William Bennet's denunciation of the schools as the worst public schools in the nation, Chicago Mayor Harold Washington launched an Education Summit in 1988 to reform the city's schools. Parents played a powerful role in the summit, as did the business community; much less in evidence were bureaucrats from the central office, university faculty, or the Chicago Teachers Union. Although Washington died early on in the reform process, its momentum continued and eventually resulted in Public Act 85-1418, which radically reshaped the contours of public schooling in the city. The new law created a local school council (LSC) in each school made up of six elected parents, two elected members from the community, two teachers, and the principal. Each LSC had the power to hire and fire principals, to control the discretionary budget (including the often substantial Chapter I and bilingual education funds), and to design new kinds of pedagogy, curriculum, and assessment to best serve children in their given community. Although dominated by parents, the LSCs also extended power to teachers by giving them a say in the hiring and firing of principals—power that traditionally had been denied them.

Much has been rumored about the success of Chicago school reform since 1988. As must be expected, any effort to reallocate power from the central bureaucracy to local schools and communities will create a tremendous amount of conflict at the local level as community members battle with one another over the proper choice of principal, the best allocation of funds, and optimal ways to educate their children. The tremendous value of *Charting Chicago School Reform* is its service in not only

documenting the inevitable conflicts but in also describing and analyzing the actual pedagogical and curricular consequences that ensued from them. Combining massive data findings generated by a sophisticated blend of qualitative and quantitative research, Tony Bryk and his colleagues allow us to perceive the spasmodic, conflicted, but ultimately hopeful workings of Chicago school reform in all of their complexity. The authors provide numerous tables throughout the text that track attitudinal changes among teachers, parents, and principals toward their schools as well as data pertaining to the allocation of school funds, school links with external organizations, and rates of faculty hires in actively reforming schools. The quantitative findings are interspersed with richly described vignettes that capture both the promising achievements of reform as well as those teachers and principals who resisted the overall spirit of civic engagement that animated the Chicago school reform effort.

The greatest contribution of *Charting Chicago School Reform* lies in its unusually sensitive and nuanced account of a major city's educational reform process. The authors make it clear that PA 85-1418 worked in the sense that it involved hitherto marginalized (and yet key) constituents, such as parents, in the daily work of public schools and that it created new possibilities for teachers and principals to focus their work not on district or state guidelines but rather on the instruction of the children in their care. At the same time, the authors abstain from romanticizing the new democratic localism in the city's schools; their ethnographic findings indicate that only a third of the public schools in Chicago really took advantage of the substantial new freedoms that PA 85-1418 made possible. For many scholars and reformers, this is an intolerably low percentage. Yet, because Bryk et al. understand that real school reform— that kind which really focuses on children and enhancing their learning— requires a sweeping transformation of the ways that most teachers instruct, they are less hurried to condemn the Chicago effort. Instead, they cite a well-established body of research that conclusively demonstrates that real changes in the instruction of teachers come not through workshops or other kinds of instruction outside of the classroom but rather through a kind of assisted performance, sustained on a regular basis over many years, in which teachers learn from more capable peers about more effective and versatile teaching repertoires. This kind of school reform evades the fads that come and go with startling rapidity in educational circles and instead requires a patient, ongoing, and daily commitment to work in classrooms with students and teachers.

There are some shortcomings to *Charting Chicago School Reform*. To my mind, the authors were too quick to dismiss the role that community-based organizations play in school reform, characterizing their work in too schematic a fashion as adversarial. Strikingly absent from the text are data measuring gains in student learning, a shortcoming that will doubtless lead skeptics to dismiss the favorable interpretation of PA 85-1418 that the authors advance. Readers should note that *Charting Chicago School Reform* does not analyze data from beyond 1995, so that

more recent, centralizing influences in the city's politics of education do not form part of the narrative or analysis. Finally, instructors interested in using the book in classes might wish to know that Bryk et al. declined to speculate on the meaning of Chicago reforms for other cities in the United States. While such circumspection reflects the overall tone of empiricism and prudence that characterizes their work, it diminishes the book's significance for a national audience.

In spite of these problems, the scholarship that animates *Charting Chicago School Reform* is breathtaking in its sweep and incisiveness and sets entirely new standards for the sociological literature on school reform. No other study on recent school reform even approaches the authors' comprehensive analysis of the transformation of a major metropolitan public school system. In addition, no other recent account of urban school reform makes such a powerful argument for the efficacy of democratic localist school politics. If the authors are accurate in their analysis and democratic localism can break through the power of an urban school bureaucracy as powerful as that in Chicago to reconnect poor and working-class parents to their children's schools, the lessons from this experience for the rest of the country could be profound and could provide a serious alternative to the current dominance of marketplace models of school reform in policy circles and the media.

*Ghetto Schooling: A Political Economy of Urban Educational Reform.* By Jean Anyon. New York: Teachers College Press. Pp. xix+217.

John A. Kovach
*Kutztown University*

Several years ago when I was working for the National Center for Education in the Inner Cities, I had a brief conversation with a grants officer at the Department of Education's Office of Education Research and Improvement (OERI). I was a bit stunned when the OERI representative, in discussing one of my center projects, said to me, "Sure, the results of this project look good, but, in the end, it's sad to know that none of these initiatives will truly make very much difference for these kids." I was shocked because this individual, responsible for improving the life chances of urban youth, clearly appeared to be resigned to the fact that these children were doomed to failure. It is precisely this personal resignation among policy makers, and the cynicism that many hold for reform efforts in urban schools, that make Jean Anyon's *Ghetto Schooling* a critically important book for all individuals interested in school reform and educational improvement. Anyon clearly illustrates how the same factors that doom urban children to failure in school and destroy the hopes of their families also smother the aspirations of teachers, administrators, and policy makers. Through its deep structural diagnosis of urban educational problems, this book makes a vitally important contribution to un-

derstanding how urban schools have ended up in such dire straits. At the same time, she prescribes specific economic and political structural changes that are imperative if reforms at the school level are to be successful.

For four years, beginning in 1989, Anyon was involved in presenting workshops for educational restructuring in the Newark, New Jersey, school district. She presents excerpts from her field notes to put a "face" on her discussion of the effects of social class and race on educational reform. These excerpts work well in giving the reader a heartfelt sense of the horrendous conditions in urban schools as well as conveying the urgent need for a solution.

Anyon's work is well researched, and she expertly applies a historical analysis to link the current crisis in Newark's schools to past political and economic decisions, trends, the influence of crime, and decades of corruption in local government. Unlike many current analyses of the failure of urban schools that blame school personnel, teachers, parents, or even the students themselves, Anyon's aim is to expose the cumulative effects of economic and political decisions in the larger urban context as they have had a devastating impact on urban schools. She clearly explains how these past decisions and current economic and political arrangements sabotage any effort to achieve meaningful school reform. This research is highly readable. Anyon's use of ethnographic notes vividly illustrates the effects that increasing poverty and social isolation of urban neighborhoods have had on educational performance of urban youth. Her crisp writing style and use of a historical approach make this book ideal as an ancillary text for social problems courses.

Finally, Anyon presents a strategy for change that reflects the lessons from the past. Her vision includes political and economic policy changes for directing a portion of national resources to reconstruction of the central city environment. What is needed for successful educational reform is "economic reprioritization and resuscitation of the cities themselves" (p. xvi). All of her suggestions for change flow logically from her diagnosis and analysis. Most important, her recommendations for change are not broad and sweeping generalizations, rather, they are concrete and quite feasible. Anyon even outlines a curriculum for teacher training, right down to the detail of suggested readings, done with such precision that one could instantly design a course syllabus from her presentation.

Ultimately, Anyon's prescriptions for change are deep and structural, which is precisely why they probably will never find their way into the current movement for urban school reform. Over 30 years of research has consistently demonstrated that academic achievement is correlated with SES; it seems obvious that educational policy must address these economic and political factors. However, to address these macroecological factors requires criticism of the status quo; it requires that corporations as well as national political and economic leaders be held accountable for decisions that result in the devastation of urban areas.

Anyon optimistically ends her book by stating that every 20–30 years during this century Americans have responded to social distress and unrestrained corporate expansion and profiteering with egalitarian reforms. She feels that this country is, perhaps, at such a conjuncture. If the United States had a strong labor party or some structure for organizing disparate groups dedicated to economic and political structural reforms, and if educational professionals were organized and linked with such a national organization, then I would share Anyon's optimism. At this point, however, the resignation of the OERI grants officer mentioned earlier seems destined to prevail for a while longer.

*From School to Work: A Comparative Study of Educational Qualifications and Occupational Destinations.* Edited by Yossi Shavit and Walter Müller, with the editorial assistance of Clare Tame. New York: Clarendon Press, 1998. Pp. xxi+550. $135.00.

Alan C. Kerckhoff
*Duke University*

This is a very important book. It brings together many of the very best social stratification analysts to conduct parallel studies of a core stratification linkage. The result is a rich scholarly feast that is both highly informative and a challenge to future stratification research.

The book is designed to examine the interface between education and entry into the labor force in 13 countries: Australia, Britain, France, Germany, Ireland, Israel, Italy, Japan, the Netherlands, Sweden, Switzerland, Taiwan, and the United States. Most of the chapters include unique features, but parallel analyses are conducted in the following ways: (1) Ordinary least square estimates are made of the association between educational attainment and the occupational status or prestige of first jobs. (2) Multinomial logit analyses are made of the odds of entering the labor force in different occupational classes. (3) Multinomial logit analyses are made of the odds of being employed, unemployed, or not in the labor force.

The chapters also describe the nature of the countries' educational systems, both historically and currently. The great variety of the structures of the systems and the kinds of credentials they award is clearly indicated. In an effort to provide a common basis for the studies, however, the analyses all use the CASMIN educational attainment categories in both the OLS and the multinomial logit analyses. The Erikson-Goldthorpe-Portocarero occupational class categories are used in the logit analyses.

Most of the chapters provide additional insight into issues of importance in the particular country or of special interest to the authors. For instance, the chapter on Israel (by Vered Kraus, Yossi Shavit, and Meir Yaish) analyzes ethnic differences; the two chapters on Japan (by Hiroshi

Ishida and by Takehiko Kariya) emphasize the importance of firm size; and the chapter on Sweden (by Robert Erikson and Jan O. Jonsson) takes particular note of the effects of gender and industrial sector.

This brief review cannot say much about the studies of individual countries except that they are consistently of high quality. However, Walter Müller and Yossi Shavit provide an analytic overview of the studies in an introductory chapter. They build on earlier research on educational structures to classify the systems according to their degree of standardization and stratification, whether they differentiate between academic and vocational credentials, and the proportion who obtain postsecondary qualifications. They generate a series of hypotheses about how these differences should affect the analyses reported.

Their comparative analyses suggest that the strength of the association between educational and occupational attainment is greater in educational systems that are stratified and standardized, that make clear academic-vocational distinctions, and that have smaller proportions with postsecondary credentials. Their analyses also suggest that entry into the skilled working class is much more predictable in societies that make an academic-vocational distinction among educational credentials. They find a general tendency for women's labor force participation to be higher with higher educational credentials and for unemployment to be greater for both men and women who have lower credentials. The combination of the detailed country studies and this integrative overview is very valuable.

I said earlier that the book is both highly informative and a challenge to future stratification research. It poses a challenge in at least three areas:

1. *Data adequacy.*—Although the focus is on the entry into the labor force, the data sets used in the analyses of Ireland, the Netherlands, and Sweden did not have data on first jobs. So, the jobs of relatively young workers were used as proxies.

2. *Definitions of first job.*—Even the analyses that used data on first jobs used varied definitions. For instance, the American analysis (by Richard Arum and Michael Hout) defines the first job as the one obtained after completing all full-time schooling. In contrast, the British analysis (by Anthony Heath and Sin Yi Cheung) deletes all individuals who obtained any full-time education after entering the labor force. In the chapters on both Germany (by Walter Müller, Susanne Steinmann, and Renate Ell) and Switzerland (by Marlis Buchmann and Stefan Sacci) first job is defined as the one obtained after completing an apprenticeship (which involves considerable labor force experience). There is no simple way to adjust for these different definitions, but they undoubtedly affect the results.

3. *Categories of educational attainment.*—The CASMIN categories are most applicable in countries in which there is a clear academic-vocational differentiation. Several of the analysts report difficulties using them, and some of the categories do not exist in some countries. I cannot tell how

much the findings might be affected by these problems, but further study is warranted.

These are all challenges worth striving to meet. The work reported in this book takes us a long way toward understanding the interface between education and labor force entry in comparative terms. It provides a firm base on which to build.

It is a pity that the book has such a high price. That will prevent many scholars from adding it to their personal collections. However, they should make sure that their university libraries obtain at least one copy. No serious scholarly library should be without this book.

*The Black-White Test Score Gap.* Edited by Christopher Jencks and Meredith Phillips. Washington, D.C.: Brookings Institution Press, 1998. Pp. x+523. $49.95 (cloth); $18.95 (paper).

George Farkas
*University of Texas at Dallas*

In 1903, W. E. B. Du Bois forecast that "the problem of the color line [will be] the problem of the twentieth century." This forecast came true, and race and ethnic inequality has been one of the principal foci of the field of sociology. Consequently, one might suppose that by 1999, sociology would have produced a body of detailed evidence on the causes of race and ethnic inequality, which could be applied to improve opportunities for race and ethnic minority children and adults.

Unfortunately, this has not occurred. Instead, when Oscar Lewis wrote about the culture of poverty, Daniel Patrick Moynihan highlighted the decline of the African-American family, and James Coleman examined the failures of busing and the public schools, they and their work were attacked amid charges of racism and "blaming the victim." The chilling effect on research in this area has been so great that in his review of *The Bell Curve* in these pages (*AJS* 101 [3]: 747–53), Douglas Massey said: "The discipline of sociology has a lot to answer for, and one of the things I lay at its feet is *The Bell Curve*. If sociologists had been more forthright in studying human intelligence [more accurately, *cognitive skill development*] over the past two decades, Herrnstein and Murray might never have written this book." Massey concludes his review by challenging sociologists to begin examining these issues.

Now, Christopher Jencks and Meredith Phillips have undertaken this task. *The Black-White Test Score Gap* includes 15 papers they commissioned for a series of workshops and meetings to assess what we know about cognitive skills and schooling in the wake of the publication of *The Bell Curve*. These papers offer two especially surprising conclusions: "First, the test score gap between blacks and whites turned out to play a much larger role in explaining racial disparities in educational attainment and income than many had realized. Second, many common expla-

nations for the test score gap, including racial disparities in family income, school resources, and innate ability, did not seem to be as important as their proponents claimed" (p. vi).

The papers in this volume provide much detailed evidence concerning these and related claims. Evidence regarding each of the following is presented and critically evaluated: the current magnitude of the black-white test score gap, as it varies from infancy through twelfth grade; the role of this gap in accounting for black-white differences in education, income, and other outcomes; the over-time trend in this gap; and the likely relative importance of each of the following in explaining this gap—test bias, school resources, genetic endowment, family background, and cultural and psychological forces and mechanisms.

The broad sweep and intricate detail of this material cannot be adequately summarized in a brief review. Instead, I will describe what appears to be most new and striking in the author's conclusions. I rely heavily on the excellent introduction and chapter 7, "Does the Black-White Test Score Gap Widen after Children Enter School?" by Phillips, Crouse, and Ralph.

When they begin first grade, black children's vocabulary, reading, and mathematics related skills and knowledge are well below those of whites. For example, the average black first grader scores below the twenty-fifth percentile of the white first-grade vocabulary distribution, a gap which is only modestly reduced by controls for family socioeconomic background. This is almost one standard deviation below the white mean; in curricular terms, the average black first grader scores at the same level as the average white kindergartner. By twelfth grade this gap has increased slightly—the average black student is scoring significantly below the twenty-fifth percentile of the white distribution. Now, however, because the standard deviation of the distributions is larger, this is equivalent to the average score for a white eighth grader. (Also see Lawrence Stedman, "An Assessment of the Contemporary Debate over U.S. Achievement," in *Brookings Papers on Education Policy 1998*, edited by D. Ravitch). Further, the black-white twelfth-grade gap is only modestly reduced by controls for family socioeconomics and for schools attended.

The authors report that more than half the twelfth-grade racial gap would be eliminated if we could eliminate the differences that exist before children enter first grade. But they also confess ignorance as to the source of the racial gap that develops during the child's first six years of life, as well as the reason that black children learn less than whites after they begin formal schooling. Jencks and Phillips conclude that "if racial equality is America's goal, reducing the black-white test score gap would probably do more to promote this goal than any other strategy that commands broad political support" (pp. 3–4).

Whether you agree with this statement or not, if you care, either professionally or personally about racial equality, you must read this book. Yet this volume is only a beginning, since as the authors note, we still know very little about *why* black children learn less than whites during the

preschool and school years. Much work will be required to understand and to change this situation.

*Social Learning and Social Structure: A General Theory of Crime and Deviance.* By Ronald L. Akers. Boston: Northeastern University Press, 1998. Pp. xx+420. $55.00.

Michael R. Gottfredson
*University of Arizona*

Ronald Akers has long been criminology's most influential advocate of social learning theory. In *Social Learning and Social Structure* he recounts the history of his involvement with the theory, describes the relationship between his learning theory and Sutherland's differential association theory, reviews contemporary research on crime and delinquency (including some of his own work) in the context of social learning theory, offers criticism of other theories, and concludes with an essay demonstrating how social learning theory may be thought of as a structural theory of crime and deviance. In so doing, Akers has written an important book for criminology. He presents a bold vision for his theory, defends it against its critics and, in the process, sweeps other social theories into the open arms of the social learning perspective. Many sociologists will appreciate greatly his efforts to link principles of social learning theory to variations in the social structure, his explicit connections between criminological and deviance literatures, and his insistence on positivistic principles for the assessment of the worth of crime theories.

Akers argues that, properly understood, theories (and the facts that support them) as diverse as self-control and strain, social disorganization, and social control fall easily within the scope and principles of the social learning perspective. It is difficult to conceive of a more thorough, more aggressive, or more articulate presentation of one of the most influential theories of modern criminology. Akers is a distinguished researcher who has long contributed significant, well-designed and well-executed studies of delinquency and deviance that provide empirical support for his views (many of which are discussed in this volume). Much of the book also continues and elaborates on the debate between advocates of social learning approaches and devotees of other theories (especially, it seems fair to say, devotees of control theory). It is, without question, an extremely important contribution to criminology. The arguments and data in the book are not likely to win over to social learning theory those who value making explicit the varying assumptions about human nature implicit in different theories of human action (or, indeed, that such assumptions are implicit in these theories). Nor will it satisfy those who believe that theories in this field should begin with an idea of crime and deviance, an idea that derives from the theory. The ready acceptance of dependent variables created by law or by discipline is in this view the fatal flaw of

positivistic approaches such as that provided by Akers. Akers's treatment of some important issues in criminology is difficult to comprehend. Although he seems to stipulate the general invariance of the age-crime curve (p. 343), he uses categories of his theory to explain this constant. He interprets the stability effect to mean that levels of criminal and delinquent activity do not change with age, rather than to the general stability of individual differences over extended periods of time (p. 351), and thereby distorts the view of rival theories. And, the now persistent finding that individual differences in the tendency to engage in troublesome behavior can be identified very early in life (well before association with delinquent peers) is not adequately confronted. Indeed, early onset, stability of individual differences despite social circumstance, and the ubiquity of the age effect are each assaults on a theory that predicts that delinquent teenage friends, changes in reference group with age, and the like cause crime.

Akers sees no value in distinguishing between theories that stress how people learn to conform their own self-interested behavior to the expectations and requirements of others from theories that stress how people learn to embrace the values, techniques, and rationalizations that require them to engage in crime in order to receive interpersonal rewards. So, he feels comfortable assuming that control theory is simply one version of his general learning theory. He sees no value in distinguishing between theories that assume that people commit crime because their conventional aspirations are frustrated from theories that assume that aspirations cannot be so classified. So, strain theory too is simply a version of general learning theory. So be it. But perhaps control theorists will be forgiven if they cling nevertheless to the view that crime represents the antithesis of culture and that it necessarily spoils rather than enhances interpersonal relations (especially within the family). Given the current status of strain theory, its proponents are probably happy to find such a receptive home.

*Ethnography at the Edge: Crime, Deviance, and Field Research.* Edited by Jeff Ferrell and Mark S. Hamm. Boston: Northeastern University Press, 1998. Pp. xviii+309. $55.00 (cloth); $22.50 (paper).

Sudhir Venkatesh
*Harvard University*

Ethnographers have studied criminal activities for decades. Yet, apart from an appendix or an autobiographical essay, they spend relatively little time analyzing the methods of their work. *Ethnography at the Edge* is unique in this regard. The 11 authors who contribute to this volume all share a deep concern for the nature of criminal inquiry including the thorny ethical and pragmatic challenges researchers of crime face. Their argument, which is a provocative one, is that a true ethnographic study of crime (a "true confession") involves some type of complicit behavior, often a performance of the very transgression one is studying. To dismiss

the moral, theoretical, and methodological conundrums that follow is to risk a very partial understanding of the social practices being studied.

Each of the chapters focuses on a particular criminal activity, albeit "crime" becomes an inclusive category since it includes homelessness, phone sex, and gang activity. Each author has faced compromising situations that have forced her or him to make difficult, on-the-spot decisions to ensure personal safety, the safety of their informants, and the viability of their respective ethnographic projects. Consuming drugs to bridge social distance with informants, being arrested in a crackhouse, observing assault and domestic violence, and recording the transactions of marijuana dealers appear to be the de rigeur initiation rites of the ethnographer of crime. These activities bring problems not only in the field but also in the reporting of data and the maintenance of fieldnotes.

The guiding thread that ties the chapters together is "criminological verstehen," in which the researcher attempts to relate the subjective components of crime, that is, situational meanings and emotions for the actors involved, to the broader contexts in which the activity takes place. From the outset, a rigidly objective research posture is denounced: not only are the standards of mainstream science impossible to realize in this type of work, but to attempt science risks losing the very fruit of one's labor, namely the criminal's perspective and the researcher's constitutive role in the research process.

The strength of the work is the introduction and the concluding chapters that analyze, in different ways but with equal care, the practice of ethnographic research into crime. They are both blunt discussions of the advantages of "true," that is, partially complicit ethnography. They make strong arguments for the need to support this work, which they see as impinged upon by institutional review boards, quantitative social scientists, and funding agencies. Unfortunately, the case studies never fully develop the analytic themes that the editors outline. There are captivating stories of heroism, bravado, and historic accidents that made fieldwork successful, but what all this implies for an understanding of the ethnographic practice is underexplored. The gulf between theory, reflection, and narrative is never really bridged.

James Clifford and George Marcus have made the point forcefully that ethnographic work does not end when the researcher leaves the field; it continues into the process of writing, theory selection, and argumentation. Because of the general conflation of ethnography with fieldwork, the case studies do not explore the politics of representation. They offer little insight into the difficulties involved in writing stories of crime for a public audience. They are more successful in using personal narrative to lay out the ethical and methodological terrain faced by the ethnographer of crime, which is no small feat to be sure. However, each chapter could have gone further by exploring the pitfalls of a reflexive sociology and the choice one makes when objectivity is given up for a more textured picture of a constructed reality.

The most engaging essay in the volume is "Marijuana Subcultures," in

which Ralph A. Weisheit examines the bias among social scientists study-ing urban crime. He argues persuasively that the study of urban crime offers a poor set of analytic tools for social transgression in the rural con-text. His study exemplifies a truly critical, reflexive sociology by ground-ing the structures of knowledge in social structure: the categories used to study crime reflect a particular social perspective, what he calls "urban ethnocentrism."

These criticisms do not detract from the innovative quality of the proj-ect as a whole. Although the constituent essays vary in terms of their theoretical and analytic sophistication, each contributes by offering an honest assesment of the experience of studying life at the margins. To-gether, they make a compelling argument for the need to redirect crimi-nology to reflexive as opposed to scientific sociology.

*Mastered by the Clock: Time, Slavery, and Freedom in the American South.* By Mark M. Smith. Chapel Hill: University of North Carolina Press, 1997. Pp. xx+303. $45.00 (cloth); $16.95 (paper).

Tukufu Zuberi
*University of Pennsylvania*

A dynamic tension existed between agriculture and industry in the early development of the modern market economy. The world demand for Bra-zilian coffee, Cuban sugar, and cotton from the southern United States resulted in the dominance of the large plantation. Slavery became an in-dustrial system, and time consciousness and clock-regulated work were intimately bound up with the new economic developments in the south-ern United States. Productivity among the enslaved increased substan-tially after 1830 (Eltis, *Economic Growth and the Ending of the Transat-lantic Slave Trade* [Oxford, 1987], pp. 286–87). In *Mastered by the Clock: Time, Slavery, and Freedom in the American South* Mark M. Smith ex-amines an element of the labor-productivity improvements that may have contributed to the physical output per slave in the 19th-century United States. Smith documents how clock time was introduced in the South and its impact on the relationship between the planter class and the enslaved Africans.

"Clock consciousness" for large-scale plantation production in the southern United States became important around 1830. Between 1830 and 1865, the use of clock time to monitor and evaluate labor time was introduced into the southern plantation. The South was one of the few agricultural societies in the 19th century to develop clock conscious-ness.

Above all else, southern planters desired to make a profit from the labor of the enslaved Africans. How did the enslaving southern planters maximize their profit, preserve racial enslavement, and pursue moder-nity? This is the question addressed by the author. Smith argues that the

planter class's adoption of clock consciousness satisfied their drive for more profit, for domination (discipline and order), and for modernization.

Smith misses the opportunity to investigate how the introduction of clock consciousness was related to the interconnectedness of the northern industrialization and agricultural production. After the 1830s, agricultural production in the South increased. This increase in productivity required more efficient exploitation of enslaved Africans. Both the domestic and foreign industrialists needed a productive South. The cotton being produced in the southern United States was an important part of industrial development in the West. The clock could empower the master with enhanced control over enslaved labor. The African resistance to clock-regulated work time reflected a challenge to their enslavement. The very fact that masters had to force obedience to clock time was evidence of the African resistance and also an indication that the masters saw this challenge as a threat to their continued dominance.

As Smith notes, "industrial capitalists and antebellum slave owners sometimes agreed: both harbored, to varying degrees, a distrust and suspicion of workers, slave and free" (p. 6). Thus, both sought to hold a monopoly over time and to control how time was defined. The debate over the use of clock time among wage laborers introduced the language and worth of clock time and brought the "working class under the ideological umbrella of industrial capitalism" (p. 7). The debate over time legitimated the industrialists desire to introduce clock time as a measurement of labor. "Time was money." Along with the importation of clocks and watches, the planter class also got the idea that time was money and that the clock could be used to regulate, measure, and exploit African labor in a more efficient manner.

The antebellum southern masters "owned" the enslaved; therefore, they owned the time of the enslaved. The slaveholders wanted slaves to use clock time to be more efficient in their productivity especially during harvest time. The wage laborers would become "clock disciplined," while the enslaved could only be "clock obedient." In the industrial north, the battle was over the idea of time; in the south, the master owned the technology and wanted the enslaved to submit to "his" time. This is why the ownership of a timepiece was not an important prerequisite for the internalization of time discipline in the South. Smith argues that the antebellum slave owners were not capitalist in the free wage-labor sense, but they became modern in their understanding of clock time. By the Civil War, the planter class adopted an obedience to clock time like their northern brethren.

The antebellum slave owners embraced the bourgeois practice of clock-regulated labor; however, they rejected the norms of freedom for the laborers. The difference was wage and freedom. The slave worked from "sun up to sun down," and this regime was not transformed by the introduction of the clock. Smith recognizes that the enslaved were not expected to develop an internalized respect for time, they were forced to work according to the dictates of their enslaver. The clock was added to

the mechanisms of oppression, such as the whip, invoked when masters needed the enslaved to labor.

The major shortcoming of this fine book is its inability to capture how the introduction of clock time in the antebellum South was another effort to transform the enslaved into manipulable objects. This process of objectification promotes the notion that people's actions have no impact on the overbearing world, the enslaved are presented as spectators not participants in constructing and reconstructing their world. Even when Smith observes actors rejecting their objectification, his theoretical orientation leads him to a dismissive conclusion. This is true in his analysis of the enslaved (p. 152) and the slave master (p. 176). The efforts by the enslaved to preserve their agency and increase the scope of their freedom are seen in the context of "identifiable, sometimes impenetrable, limits" (p. 136).

*Russia's Economy of Favours:* Blat, *Networking and Informal Exchange.* By Alena V. Ledeneva. New York: Cambridge University Press, 1998. Pp. xiii+235. $59.95 (cloth); $21.95 (paper).

Federico Varese
*Nuffield College, University of Oxford*

The Soviet Union outlawed market mechanisms for the distribution of most of the goods and services it produced. It relied instead upon a number of criteria for the distribution of goods, benefits, and burdens: "first come, first served," urgent need, merit, family background, social status, and political conduct. The Communist leadership believed that, if properly implemented, such criteria would make Soviet society more fair than its capitalist counterparts. Also, it was forced to ration many consumer goods because of productive inefficiency: demand was higher than supply. The Soviet state invested a great deal of resources in an attempt to make this system of allocation work. Yet at the same time, Soviet citizens actively subverted the system to a degree not attainable by any foreign power, nor by any of the few, internal dissidents. They did so by using *blat,* the subject of Alena V. Ledeneva's fascinating ethnography of a crucial yet underexplored aspect of Soviet society.

Ledeneva reminds the reader that *blat* is an elusive social practice; "*blat* is an obvious word, it needs no definition" says one of her respondents. Every Soviet citizen knew what *blat* could bring him or her: foodstuffs and consumer goods, adequate health care, housing and household upkeep, theatre tickets, employment, job transfers, and job promotions. Parents would anxiously search for *blat* in order to obtain a place in nursery school, high school, and university for their children or to relieve them from military service. *Blat* was an extensive network of connections and at the same time a reciprocal relationship governed by the principle *Ty-mne, ya-tebe* (you help me, I help you).

Ledeneva skillfully conveys the rich texture of this social practice and the strategies deployed by ordinary citizens. She draws on 56 in-depth interviews carried out mainly in the period 1994–95, literary sources, archival material, and a review of cartoons and satirical articles that appeared in the journal *Krokodil* from the 1930s to the late 1980s. Where possible, Ledeneva supplements her ethnography with survey data, although the absence of longitudinal studies that go back to the Soviet period limited her use of quantitative data. This book makes a substantial contribution to the sociology of everyday life in the Soviet Union and to the way ordinary people circumvented the socialist system of distribution. In the final chapter, Ledeneva takes her subject to the post-Soviet period by considering the role of *blat* in the new market economy.

At the end of a chapter tellingly entitled "*Blat:* The Unknown Common Place," Ledeneva defines her object of study as "an exchange of favours of access in conditions of shortages and a state system of privileges. . . . *Blat* is a distinctive form of non-monetary exchange, a kind of barter based on personal relationship" (pp. 34–37). *Blat* favors were carried out informally through networks based on "mutual sympathy and trust" and implied delayed reciprocity. Such favors were provided at the expense of the general public and served the needs of personal consumption. In a different part of the book, Ledeneva explores the social rhetoric of friendship and acquaintance that surrounded *blat:* "sharing," "helping out," "friendly support," and "mutual care" all were used to characterize a practice that was illegal and socially disapproved yet at the same time unavoidable and widespread. Most respondents engaged in what Ledeneva calls "strategies of misrecognition" (from Bourdieu): one could see *blat* in others but would (at least initially) define his or her own identical behavior as acts of generosity or altruism.

The range of situations in which *blat* was used in the Soviet Union, the author says, renders it irreducible to some other concept: *Blat* is similar to but also different than "second economy," "black market," "political corruption," "bribery," "patronage," and "neotraditionalism." This is a bold line to take. Most authors would argue that the word *blat* was used to refer to analytically different behaviors and would try to reduce the practices described as *blat* to existing concepts. In my view, for instance, some cases of *blat* could be reduced to barter exchange among trusted partners and many other cases to corruption: an "agent" (an official), instead of implementing the rules laid out by his "principal" (the Soviet State), bent them in order to favor a "corrupter" for something in return. When trust between the official and the corrupter was lacking, the corrupter sought a mediator. Mediators lessened the risk of defection by officials and were compensated. The peculiarity of the Soviet Union lay in the spread of such behaviors, making corruption the norm rather than the exception. Having *blat* meant that one could carry out corrupt exchanges more effectively than those without it. Money was rarely used because the rouble had little purchasing power in the planned economy.

The "antireductionist" line taken by Ledeneva may nevertheless prove

to be the best one. It allows her to explore in full the complexities of *blat* and to produce a book that is a pleasure to read. *Russia's Economy of Favours* will surely remain the standard reference on this subject for years to come and hopefully will initiate a virtuous exchange of ideas—not favors—between sociologists, anthropologists, economists, and social historians.

*Informal Politics: Street Vendors and the State in Mexico City.* By John C. Cross. Stanford, Calif.: Stanford University Press, 1998. Pp. ixx+272. $49.50 (cloth); $18.95 (paper).

Lauren Benton
*NJIT and Rutgers University, Newark*

This interesting but sometimes frustrating book seeks new insights about the growth of the informal, or unregulated, economy. The author identifies a gap in the earlier literature on the phenomenon: relative inattention to the political connections between the state and informal sector actors. The book provides an often fascinating look into the politics surrounding street vending in Mexico City.

However, in responding to the omissions of earlier informal sector researchers, who often isolated economic from political analysis, the book overcorrects. It attempts to build a theory of informal sector growth based on analysis of state integration and informal sector visibility and political organization. In doing so, it largely omits factors identified in earlier literature as essential to an understanding of informal sector change, namely patterns of sectoral structuring and entrepreneurial culture

The research on street vending is carefully done and in many ways original. Cross traces the history of street vending in Mexico City and, in particular, the history of state interventions. Like other good monographs on the informal economy, *Informal Politics* is persuasive in arguing that history matters; it shapes the strategies of both state and informal sector actors in the present. The book is also particularly valuable for its careful study of the political organization of street vendors. Cross shows that the state's attempts to control street vending ultimately prompted the creation of informal political organizations and that these in turn made the state's efforts at control fail. In this sense, the central argument—that gaps in state regulation are the result of neither calculated functionality nor state inefficiencies—is wholly convincing.

However, in order to set up this argument, the author provides what is ultimately a caricature of earlier studies of the informal sector. He sets Hernan DeSoto's view that the informal sector represents untapped economic potential, which can be captured through deregulation, against a "neo-Marxist" school, which sees the informal economy as subordinate to, and exploited by, the formal sector. This contrast is misleading. As

Cross himself admits (p. 27), the "neo-Marxist" school, though it began with an insight about the benefits of informal labor for formal sector firms, shifted to portray informality as a phenomenon with different meanings and impacts on workers in different political and cultural settings. What Cross seems to miss is that this shift was partly prompted by a dialogue between informal sector analysts interested in the developing world and scholars tracking rapid sectoral restructuring, including informalization, in advanced economies. The dialogue revealed a wider range of outcomes of informalization and led to calls for a more detailed study of the political and cultural factors producing this complexity (A. Portes, M. Castells, and L. Benton [eds.], *The Informal Economy: Studies in Advanced and Less Developed Countries* [Johns Hopkins University Press, 1989]).

Cross's study fits squarely into this project and is a valuable contribution. But since he sees his approach as separate, he misses opportunities to draw on the insights of earlier research in order to ask some crucial questions. Do social networks among street vendors shape their participation in political organizations? Do patterns of sharing resources create patterns of solidarity that affect both political and economic actions? What are the vendors' relations to distributors, and how does their insertion in the wholesale-retail chain more broadly respond to, or help to determine, their political strategies? What is the politics of gender in this sector, where many of the "autocratic" leaders of vendor organizations are women? Without attention to such questions, Cross must limit his analysis to the impact of informal politics on informal sector "growth" (p. 241) rather than producing an understanding of the kind of growth likely to occur.

Some of my frustrations with the book no doubt stem from problems of presentation. We do not find out how many street vendors are estimated to be operating in Mexico City until page 86. The book does not address the relation of street vendors and the state in Mexico—its main subject—until chapter 5 (p. 120). The writing, though clear, would have benefited from more careful editing (to give just two examples, "on the other hand" is rarely paired; Colombia twice becomes "Columbia").

In summary, the book confirms that state structure is important to an understanding of the informal economy. It uses original data to show that informal-sector political associations are influenced by, and in turn influence, state policies toward informality. The data on street vending will be of interest to scholars of urban Mexico, and the discussion of "informal politics" is a useful supplement to earlier approaches and deserves notice by development scholars, economic sociologists, and state theorists.

*Territorial Ambitions and the Gardens of Versailles.* By Chandra Mukerji. New York: Cambridge University Press, 1997. Pp. xxiv+393. $79.95 (cloth); $34.95 (paper).

David A. Bell
*Johns Hopkins University*

Visitors to the gardens of Versailles, no matter how aware of the horrors that King Louis XIV inflicted on his people in the name of glory and the Roman Catholic Church, have a hard time not being impressed. Here is an artistic transformation of nature on a breathtaking scale: earth shifted; trees, bushes, and flowers planted; canals, ponds, and fountains built; paths laid out and statues placed, acre after acre, seemingly without end, mostly to glorious effect. One immediately senses that whatever else Louis XIV intended to accomplish here, he certainly meant to convey one very simple point: only a great king could do this.

But as Chandra Mukerji shows in her study of the gardens and French political culture, this is only the starting point. In a marvelously well-illustrated book, based on an impressive mastery of technical literatures on everything from fountain building to tapestry weaving, Mukerji reveals the many different ways in which the gardens served as arenas for Louis XIV's ambitious attempts to remake his state. The large-scale earthworks allowed military engineers to demonstrate and experiment with new methods of fortification. The design and ornamentation provided a showcase for artists charged with developing a new, specifically French aesthetic style. The gardens also provided an important setting for the famously elaborate ballet of power and etiquette through which the King maintained control over the high aristocracy. In perhaps her most original chapter, Mukerji suggests that the gardens of Versailles can be seen as part of a process by which early modern urban elites reclaimed and reshaped the countryside, applying techniques first developed in Paris and other cities to the transformation of rural settings.

Mukerji also makes a more ambitious argument, but it is less convincing since she presents little evidence beyond a reading of the gardens themselves. Briefly, she suggests that the massive reshaping of land in the gardens of Versailles served as a laboratory for something Louis XIV and his ministers hoped to carry out on a much larger scale: the wholesale transformation, both institutional and physical, of France itself into an orderly, well-bounded "territorial state" through which royal power could flow without interruption or dilution. By symbolically joining the state to the very earth of France in the carefully shaped spaces of the gardens, she continues, the crown hoped to provide legitimation for an absolute monarchy that consisted not only of a divinely ordained king but also of a new, large, and distinctly undivine bureaucracy. Building on these ideas in her conclusion, Mukerji issues a welcome and stimulating call for social scientists to pay renewed attention to human engagement with the

material world and to the material aspects of state power, stressing the limits of historical approaches that reduce all social phenomena to texts.

For a work by a nonspecialist in French history, there are few errors, but some have crept in nonetheless (e.g., Henri IV did not reign as a Protestant; Jansenists did not simply believe in "divine grace" but in "efficacious grace"). Mukerji also fails to give readers a basic overview of the gardens, with the dates of construction, precise physical dimensions, and figures on how much it all cost. More seriously, despite her insistence on taking the material dimension of history seriously, Mukerji overestimates the material resources the French state had at hand actually to realize a new political order on the larger territory of France. The Canal du Midi was not a national highway system, and Vauban's fortress system was not the Maginot Line. Louis XIV's changes in the landscape may have pointed toward a radical reorganization of French space, but they did not accomplish it. Furthermore, they did not directly touch the lives of most of the king's subjects. While important on their own terms, in the context of producing a new France, they were largely symbolic.

The book's principal flaw, however, is its unrestrained tendency toward repetition, verbosity, and overstatement: here is a garden of learning in severe need of pruning (Mukjerji's editor did author and readers alike a disservice by not insisting on revisions). Mukerji really does not need to tell her readers that "members of cultures frequently transform both their environment and the world of creatures living around them" (p. 302) or that "our relation to nature does not begin and end with food" (p. 304). Given the recent trends in cultural studies to view all art forms as laboratories of power, it is not hugely surprising to see gardens treated in this manner, and Mukerji wastes many pages insisting on the novelty of her enterprise. The profusion of technical detail that she provides on subjects ranging from noble estate management to the international exotic plant trade to court balls, often going back in time long before the construction of Versailles, is fascinating but does not always do much to advance her argument. Moreover, much of it is far too familiar (e.g., Louis XIV's famously choreographed daily routine, discussed at length by Norbert Elias and countless historians) to need the detailed recapitulation she provides. In fact, there are really two books here: a sophisticated theoretical argument about land as a laboratory of power and a well-informed but, nonetheless, essentially antiquarian history of the gardens of Versailles. There is much to be learned from both, but it would have been better to put them under separate covers.

*Identity in Formation: The Russian-Speaking Populations in the Near Abroad.* By David D. Laitin. Ithaca, N.Y.: Cornell University Press, 1998. Pp. xiv+417. $55.00 (cloth); $22.50 (paper).

Mark R. Beissinger
*University of Wisconsin, Madison*

Dozens of books are published each year on identity politics. Rare is the work that so excels that it raises the level of academic discourse to a new plane. David Laitin's provocative study of the Russian-speaking minorities of the post-Soviet states does precisely that. It is a book of remarkable theoretical scope and empirical richness, one which stands as a model of how social scientific inquiry ought to be conducted.

Based on his tipping-game theory of linguistic behavior outlined in other works, Laitin argues that identities shift in cascade-like fashion based on what individuals believe others will do. He sketches out two types of games played within nationalist politics—language revivals and competitive assimilation—both of which emerge from the logic of the tipping-game. This book focuses on the latter of these. The core of Laitin's argument is not so much the tipping-game, but rather the choice model or payoff function, which Laitin attaches to the game, moving individuals along these curves. Linguistic tips, he argues, are the result of strategies selected by individuals that are determined by the combined effect of the expected economic returns to those adopting a language, ingroup status (the degree to which cultural entrepreneurs within the potentially assimilating group raise the status of refusing to give up one's culture), and out-group status (the degree to which members of the hegemonic culture accept those who attempt to assimilate to it).

This choice model is applied to explaining the politics underlying assimilation of non-Russian nationalities in the Soviet Union and the Russian-speaking diaspora of the post-Soviet states. Laitin portrays three patterns of incorporation of non-Russian nationalities in the USSR: the "most-favored lord" approach, providing upward mobility for those who assimilate (exemplified in Laitin's study by the Ukrainians); the "colonial" approach, with mobility opportunities partially blocked (illustrated by the example of the Kazakhs); and the "integralist" approach, in which an expanding state is unable to undermine the cultural integrity of the society that it conquers (typified in Laitin's study by the Latvians and Estonians). He then goes on to show how these patterns of incorporation have influenced the ways in which Russians in the non-Russian republics of the Soviet Union have reacted to pressures for linguistic assimilation in the wake of what Laitin calls the "double cataclysm" that has overtaken them—first, the passage of republican language laws in 1989, which altered their previous social and political dominance, and second, the collapse of the USSR itself, which subjected them to the full authority of the nationalizing state.

Laitin employs an impressive battery of methods—ethnographic, sur-

vey, experimental, historical-sociological, and content-analytic—to probe the ways in which Russians and Russian-speakers are adapting to post-Soviet circumstances in Estonia, Latvia, Ukraine, and Kazakhstan. Several chapters relay the findings of ethnographic research conducted by Laitin and three collaborators on family strategies of adaptation to state pressures for linguistic change. Laitin himself lived with an Estonian family and enrolled in an Estonian language class for Russian-speakers in the city of Narva in 1993–94 to understand better the individual-level logic causing the new Russian minority there to accept language change. Laitin's use of a matched-guise test to elicit the concealed status assumptions associated with speaking particular languages is especially ingenious. He finds that Russian speakers lose in-group respect and friendship from Russians and from titulars when speaking in their titular guise, so that there are few status incentives for Russians to assimilate linguistically (though those incentives that do exist vary considerably across his cases). This research is all the more impressive given that Laitin is not a Russian area specialist by training but rather learned Russian and gained a knowledge of this region specifically in order to conduct this project.

While the overall assumptions underlying Laitin's game theoretic analysis find support in his empirical research, he also shows that assimilation is not the only game being played and that Russians were also assessing the viability of alternative strategies of potential returns from emigration, political mobilization, incorporation without cultural change, and violent challenge. As he notes (p. 250), the opportunities and constraints set by the policies of the nationalizing state are not directly included in his model, though they are critical to the outcomes of identity politics. Moreover, in many circumstances there may be little choice actually involved in what people do, given the limited range of opportunities with which they are presented. Ultimately, as Laitin himself cautions, his choice model is not be proved by these data, but only suggested; the true test will only come after many years, when one should observe a rapid assimilation in those post-Soviet republics in which economic growth and expected payoffs for assimilation are the greatest.

One of the more controversial arguments of the book has to do with the Laitin's speculations concerning the emergence of a new identity within these beached populations, many of whom are not ethnically Russian, but Ukrainians, Belorussians, Armenians, and others. Laitin sees a "Russian-speaking" (*russkoiazychnyi*) identity emerging as an alternative to full assimilation as titulars or to mobilization as Russians. The term "Russian-speaking population" did not exist before the disintegration of the Soviet Union. It has become a significant identifier since the Soviet collapse (though, as Laitin's content analysis shows, still less frequent an identifier than "Russians"). Laitin predicts that language rather than ethnicity will eventually become the chief factor underlying identity for these minorities and argues that their ties with a Russian homeland are relatively weak, though the evidence that he presents is ambiguous. The

term "Russian-speaker" implies a strong politics of interethnic coalition and mobilization on the basis of language. This coalitional and mobilizational politics is currently not visible in any of the post-Soviet states with perhaps the exception of Ukraine, though Laitin's argument that language will become a central focus for identity hinges critically on its salience. Indeed, the extent to which Laitin reduces issues of identity to issues of language learning is problematic. Yet, it is this same simplification that allows him to pursue his game theoretic enterprise with such rigor.

Despite these issues, Laitin's book is a landmark in scholarship on nationalism and on the former Soviet Union more specifically. The boldness of its assertions, its dazzling design and execution, and the wealth of stimulating ideas found within it make it a truly outstanding achievement.

*Women in the Holocaust.* Edited by Dalia Ofer and Lenore J. Weitzman. New Haven, Conn.: Yale University Press, 1998. Pp. vii+402. $30.00.

Joan B. Wolf
*University of Wisconsin, Milwaukee*

The introduction to *Women in the Holocaust,* edited by Dalia Ofer and Lenore J. Weitzman, begins with the following question: "*Why women?* Why should a book on the Holocaust—which targeted all Jews for annihilation irrespective of their sex or age or any other social characteristics—focus on women?" (p. 1). While many social scientists might find the answer self-evident, Ofer and Weitzman's audience includes scholars devoted exclusively to the Holocaust, many of whom have been reluctant to examine questions of sex and gender. The task of this volume is thus twofold: first, to persuade those either skeptical or hostile to the inquiry that gender is a priori a useful category of analysis in the study of the Holocaust; and second, to demonstrate the ways in which gender analyses "lead us to a richer and more finely nuanced understanding of the Holocaust" (p. 1). Though, like most edited collections, *Women in the Holocaust* is uneven, stronger in certain chapters than in others, it asks new questions about and enhances scholarly perspective on the Holocaust and should be of interest both to students of the Nazi genocide and to scholars of gender more broadly defined. Whether its contributors persuade opponents of gender research on the Holocaust has more to do with debates within Holocaust studies than with the cogency of their arguments.

The text of 21 articles is divided into four sections, each of which focuses on women and gender in a particular thematic domain: Jewish life in Western and Eastern Europe between the world wars, life in the ghettos, resistance and rescue, and labor and concentration camps. It includes personal narratives of survivors, one piece of fiction by Ida Fink, and a more theoretical collection of articles at the end. Perhaps its greatest strength is to show, beyond motherhood and sexuality, a variety of ways

in which gender shaped one's experience of the Holocaust. As Paula Hyman demonstrates, for example, because Jewish women in the West were generally excluded from the world of business, higher education, and politics, they "had only limited opportunities for radical assimilation" (p. 28). Jewish women were less likely to convert or to marry non-Jews than were Jewish men, and as a result they were less likely to have Christian spouses or kin to protect them during the Holocaust. In Eastern Europe, where most Jews did not achieve middle-class status, Jewish women, of necessity, were active in secular and public life. Because more men fled their families in anticipation that the Nazis would not harass women, as shown in the selections by Marion Kaplan, Michal Unger, and Renée Poznanski, ghetto populations were disproportionately female. In this context, Eastern European women were able to draw on the resources and survival skills they had cultivated before ghettoization.

Certain contributions, on the other hand, show that gender is not always salient. In "Ordinary Women in Nazi Germany," Gisela Bock argues that gender did not determine the degree of Germans' involvement in racism and the Holocaust, that racial hierarchy prevailed over gender hierarchy, and that female perpetrators "were perpetrators not so much because they were female but because they believed themselves to be ordinary Germans, like the men" (p. 94). She concludes nonetheless that "the fact that for the Nazis race as a *political* category was more crucial than sex should not blind us to the importance of gender as an *historical* category" and that other potentially fruitful avenues of gender analysis in the study of the Holocaust remain unexplored (p. 96). Lawrence Langer argues in "Gender Suffering? Women in Holocaust Testimonies" that it is impossible to make gender distinctions in survivors' expressions of anguish. "The origins of humiliation were often dissimilar for men and women, because womanhood and manhood were threatened in various ways," he writes. "But the ultimate sense of loss unites former victims in a violated world beyond gender" (p. 362). For Langer, a gender analysis of the Holocaust experience purports to make sense of a phenomenon beyond all human capacity to comprehend. and we must "learn to face chaos with unshielded eyes" (p. 362).

Langer, in fact, implies that there is little to be learned about the Holocaust from *ever* using gender as a category of analysis, and he is not alone in this assessment. A large and outspoken number of Holocaust scholars and critics have argued that gender research is part of a broader trend toward popularizing and thus "banalizing" the Holocaust, that it diminishes the singularity of the Holocaust, and that the use of gender and many other contemporary categories of academic analysis ignores the reality that the Holocaust cannot be conceptualized with the tools applied to the rest of history. Lest this resistance to exploring the influences of gender appear to be overstated, the reader may refer in the text to "The Split between Gender and the Holocaust," where a reprinted letter from critic Cynthia Ozick accuses Joan Ringelheim, one of the first scholars to study women and the Holocaust, of "asking the morally wrong ques-

tion, a question that leads us still further down the road of eradicating Jews from history." The Holocaust, wrote Ozick, "happened to victims who were not seen as men, women, or children, but as *Jews*" (pp. 348–49).

What many of the contributions to this volume suggest is that experiences of anticipation, deportation, internment, work, and resistance were at least partially shaped by gender, or that gender can explain a particular *dimension* of Jewish suffering, just as can age, occupation, level of education, and whether one had children. As Sara R. Horowitz states in "Women in Holocaust Literature: Engendering Trauma Memory," "The bringing together of both sets of details—one particular to women, one relevant to all Jewish inhabitants of the camps—implies that survivors' reflections are inevitably gendered, and also that gender does not constitute the totality of one's experience" (pp. 370–71). Though *Women in the Holocaust* might not be able to resolve the longstanding and emotional debate within Holocaust studies, it points the way to multiple new avenues of research for those who wish to pursue them.

*The Politics of Duplicity: Controlling Reproduction in Ceausescu's Romania.* By Gail Kligman. Berkeley and Los Angeles: University of California Press, 1998. Pp. xv+358. $50.00 (cloth); $19.95 (paper).

Joanna Goven
*University of Canterbury, New Zealand*

Gail Kligman's *The Politics of Duplicity* develops two distinct but interrelated projects. The book, as Kligman describes it, is both an ethnography of the state and an ethnography of the politics of reproduction. She defines an ethnography of the state as "an analysis of rhetorical and institutionalized practices of the state within the public sphere and their integration into daily life" (p. 3). By the politics of reproduction she means "the complex relations among individual, local, national, and global interests that influence reproductive practices, public policy, and the exercise of power" (p. 5). Kligman combines the two projects by employing the lens of Ceausescu's "political demography"—a term meant to signify the extent to which the Romanian state attempted to manipulate demographic factors for political and ideological ends—to examine the processes through which state power was constituted and exercised. (Western agencies are also seen to have played a not insignificant role here.) While Kligman notes the link between Ceausescu's nationalism and his demographic policies, that is not her focus; instead, she analyzes the ways in which demographic policies became everyday techniques of control. The analysis of Ceausescu's demographic policies thus becomes a means by which "to explore the institutionalization of social practices, such as duplicity and complicity, and of identities that together constituted the Romanian socialist state and everyday life" (p. 2).

Kligman begins with a discussion of strategies of control and economic development in state socialism generally and Romania in particular. She then examines three arenas of political demography: the changing legal regime of reproduction, state provision of assistance to families with children, and the use of the medical profession to monitor the population and to enforce state policy. Following this is an analysis of the changing content of official demographic discourse and the ways in which it is productive of dissimulation. She discusses the practices through which dissimulation was institutionalized and official reality was confirmed by those whose everyday lives embodied an altogether different reality. The next chapter gives us some insight into that lived experience through interviews with Romanians, which illustrate vividly the impact of Ceausescu's coercive reproductive policies, including their corrosive effect on social relations. Finally, lest anyone think the passing of the Ceausescus has brought an end to reproductive coercion, Kligman points to the ways in which the process of transition and its associated economic inequalities have both transformed and perpetuated the suffering and violation characteristic of Ceausescu's demographic regime.

Employing the lens of political demography, Kligman identifies two key elements of Ceausescu's regime: control through homogenization and a culture of duplicity. Kligman argues that Ceausescu's neo-Stalinist project of the creation of a single legitimate identity, the socially productive "new socialist person," became a far-reaching mechanism of control. All aspects of life and personhood were implicated in this identity; thus the project denied any distinction between public and private domains and required state involvement in everything from daily diet to gynecological exams. The single legitimate identity became a stick with which to beat (arbitrarily defined) "deviants," thus creating a climate of fear and vulnerability Kligman convincingly demonstrates the ways in which the Ceausescu regime institutionalized and routinized the manipulation of people's fear to secure compliance and the affirmation of official reality.

Kligman also shows how the very intrusiveness of the state was productive of duplicity and dissimulation. The more state policies aimed to make the private transparent, the more deeply ingrained became the practices and culture of duplicity, and the more fragile became the structure built upon a population's duplicitous public selves. This duplicity both sustained the regime and eventually undermined it. Similarly, as Ceausescu's determination to pay off foreign debt generated severe domestic shortages, people were forced into increasingly desperate competition with each other for essential goods; yet this atomizing pressure was countered by the need to rely ever more heavily on informal networks. That is, the forced shortages spawned a culture of corruption, which both undermined the formal structures of control and enabled their continuation. Both the general culture of duplicity and the more particular "culture of corruption" have had destructive legacies, depleting the store of social capital for life after Ceausescu.

*The Politics of Duplicity* makes a significant contribution to the effort

within the social sciences to develop a more informed understanding of how actually existing socialism actually existed. (More work is necessary to determine which practices were characteristic of state socialism generally and which were specific to particular regimes.) It also provides compelling evidence that reproductive policies divorced from socioeconomic realities produce an appalling range and depth of physical and psychological harm.

*Languages of Labor and Gender: Female Factory Work in Germany, 1850–1914.* By Kathleen Canning. Ithaca, N.Y.: Cornell University Press, 1996. Pp. xiii+343. $42.50.

Julia Kirk Blackwelder
*Texas A&M University*

*Languages of Labor and Gender* traces the emergence of women's work in textile factories as industrialization proceeded in Germany in the 19th century. Kathleen Canning examines the labor movement, social reform efforts, public policy, and the ideology of gender. This complex group of issues is a large order for a single monograph, but Canning has succeeded in weaving these various themes into her text. The book convincingly demonstrates that gender was at the core of a transformation of the economy that reshaped the law, the state, and the family. Using a multiplicity of sources, Canning examines how male workers, social reformers, feminists, lawmakers, and bureaucrats reacted to the rapid expansion of female factory labor. While employers enter Canning's analysis as active forces of change in that they hire women both to lower labor costs and to satisfy a growing total demand for workers, the reader does not learn much about their gender attitudes.

Canning argues persuasively that male workers' definitions of women's abilities and of gender prescriptions depended heavily on the character and stability of male employment. In the Rhineland and Westphalia, where mining and metal production flourished, women replaced male textile labor more rapidly than elsewhere and male workers more willingly ceded textile employments to women. In other areas, men competed with women for factory textile employments and men invented or redefined notions of craft or skilled labor to distinguish their labor from the supposedly inferior work of women. Male textile workers also laid claim to citizenship as a male prerogative, thus striving to establish political power and simultaneously to distance themselves from female workers.

Much of Canning's text covers the contested gender ideologies of materialist labor reformers, of feminists, of politicians, and of working men. In the 1870s, responding to labor reformers who portrayed women as weak and vulnerable, Germany adopted protective legislation. Canning argues that popular views of female fragility both held women in an inferior position in the labor market and helped to unseat Otto von Bismarck,

who opposed protectionism. Protectionist views, Canning maintains, were later undercut by feminists, factory inspectors, and medical authorities who argued that women had the physical ability to labor in factories without endangering their own health or their capacity for bearing children. In either case, the issue of female employment, particularly the employment of married women, demanded action by the state in the form of protectionist codes or a right to work posture. Thus, industrialization drew the state into the private sector and into the regulation of family decisions about employment.

Canning makes a major contribution to our understanding of class relations among women in exposing the contest between middle-class maternalists and feminists in constructing images of working women that best suited their own political ends. Maternalists found women's differences from men and demanded both shorter working hours and protections to permit childbearing and domestic obligations. Feminists, in contrast, found female superiority in women's work records and in their juggling the responsibilities of the home with those of the workplace, and feminists presented their conclusions as argument for gender equality before the law and in the workplace.

Canning also set out to bring the long-gone textile women back to life, but in this effort she has largely failed. In her desire to capture the thoughts of workers, Canning turned to the types of documentation that have supported the new labor history: factory roles, vital statistics, factory inspectors' reports, and journalistic accounts. All of the above yield helpful information and allow Canning to demonstrate, for example, that the turnover rates for women and men were similar. Thus, Canning establishes that contemporary rationalizations of the gender gap in wages and the labor movement's disregard of women were wrongfully linked to arguments about the temporary nature of women's employment. On the other hand, the records reveal little about the content of the work culture that Canning has tried to extract from them. Canning draws her analysis of shop-floor behaviors from the accounts of middle-class observers, one a social reformer and another a journalist. Sporadic individual acts of rebellion or nonconformity, group celebrations, and the occasional strike are Canning's additional evidence of women's work identity. The events noted are few and far between, and the subtext is Canning's creation. Canning claims on thin citations that "rape and other forms of sexual coercion were widespread in the mills" (p. 309), but she fails to provide insight into working women's understanding of sex in the workplace and how this understanding may have shaped their attitudes toward industrial employment. In her effort to rekindle workers' voices, Canning has overreached her sources, but she has also provided a provocative perspective on factory life.

*Workers' Paradox: The Republican Origins of New Deal Labor Policy,*
*1886–1935.* By Ruth O'Brien. Chapel Hill, N.C.: University of North
Carolina Press, 1998. Pp. xiii+313. $39.95 (cloth); $17.95 (paper).

Dan Clawson
*University of Massachusetts, Amherst*

Ruth O'Brien challenges the conventional wisdom that the current sys-
tem of labor laws took shape in the 1930s. Most scholars hold that "with-
out the 1929 crash of Wall Street and the ensuing Great Depression, the
Wagner Act would not have materialized" (p. x), but she argues that "key
members of the Republican Party were responsible for developing the
fundamental principles underlying modern labor policy" (p. x) and that
they did so in the 1920s. This central claim needs to be taken seriously,
despite the book's many limitations.

During World War I, in order to gain labor peace the government of-
fered at least the moderate AFL and railroad unions a measure of protec-
tion; as a result, their membership grew exponentially. In O'Brien's ac-
count, the key issue of post–World War I labor policy was what to do
about the railroad unions. The unions proposed an extension of the war-
time policy, but this was never a serious option.

Instead, what was adopted, altered by court rulings, and then passed
in a new form was legislation that "transformed labor law" (p. 90) and
became the foundation for the Wagner Act and all subsequent U.S. labor
law. Under this system, unions as such were not represented. Instead,
workers could elect their own representatives, whether these were or were
not union, and these people would temporarily represent workers. Instead
of guaranteeing rights for unions, the legislation protected the individual
rights of *workers.* Associated with this, the new law "shifted the emphasis
of labor law from the agent to the collective bargaining process" (p. 94).
For example, current labor law requires both labor and management to
bargain in good faith, that is, to sit at the table and (at least pretend to)
listen to the other side's offer. The law does not mandate any particular
outcomes, however, and no substantive position is a priori ruled inadmis-
sible.

By the end of the book, O'Brien's evidence and text sustain a limited
version of this argument, but *Workers' Paradox* could serve as "exhibit
A" for those who argue that publishers no longer seriously edit the books
they publish. *Workers' Paradox* makes a contribution, but three of its
limitations make it challenging to read, worth the effort only for those
committed to covering all significant studies in the field.

First, basic terms are never fully or clearly explained, and what expla-
nation there is often comes a great many pages after a term is introduced.
"The principles of agency," for example, are not really explained until a
paragraph on page 186. The explanation probably needed to be at least
three paragraphs, but that is much less of a problem than the fact that
the term is central to the first 150 pages.

·.Second, O'Brien does not consistently provide the needed social context.and background. The 1919 steel strike goes essentially unmentioned, and the Seattle general strike gets passing mention only because of its direct importance to congressional hearings. More significantly, O'Brien never steps back to consider why it was that the key labor legislation of the 1920s applied only to railroads, and much of it only to the operating unions. The obvious point is that these were the unions with the greatest ability to disrupt the functioning of the capitalist system as a whole, because of the enormous importance of railroads to the economy of that time.

Third, despite—perhaps because of—O'Brien's neutral and scholarly style, she frequently uses unexamined terms and language that are far more biased than would be found in an openly polemical work. Terms like "the public good" (p. 66) and "the public interest" (p. 13) are used without quotes or irony. We learn that Republicans "restored the ideological balance" (p. 16) by passing a law that seems to me anything but balanced. In O'Brien's world, "a deep antiunion spirit had gripped the nation after World War I" (p. 96), and "a sense of urgency pervaded the nation in 1918 and 1919 when four million workers went out on strike" (p. 74) as if "the nation" were a clear and unitary construct such that no explanation or discussion were needed.

*Workers' Paradox* makes a convincing case that important elements of current labor law were shaped in part by Republican congressional legislation of the 1920s and later incorporated into the Wagner Act. This is a useful footnote to the history of labor law, but every policy has antecedents, and I am still much more struck by the Wagner Act's sharp break with previous legislation. O'Brien, a political scientist, emphasizes congressional activity and political parties; although she necessarily devotes considerable attention to court cases, she gives far too little weight to the vastly unequal social power of the competing forces.

*Philosophies of Integration: Immigration and the Idea of Citizenship in France and Britain.* By Adrian Favell. New York: St. Martin's Press, 1998. Pp. xiii+288. $65.00.

Carl-Ulrik Schierup
*Umeå University*

A proliferating academic debate and literature on citizenship has come to dominate understandings of migration ethnicity in Western Europe. Adrian Favell's comparative study on "the idea of citizenship in France and Britain" is one of the most illuminating contributions to this debate.

From the vantage point of political sociology and institutionalist theory, Favell focuses on "the philosophies of integration" belonging to these two countries. They have both had a prolonged experience with immigration, and in each a political consensus across most of the political spec-

trum has been reached, clustered around certain core ideas that are believed to promise a lasting solution to the national "ethnic dilemmas" in question. The ideas and overall institutional frameworks upon which their national policies build are starkly contrasting, however, and in each of the two nations there exists a strong public conviction that its own policy ought to be *the* model to pursue even for others.

In France, integration policy moved to the center of the political stage from the mid-1980s and was subsequently reformulated by liberal intellectuals in highly principled terms. The resulting "new republican synthesis" emphasized the cultural integrity of the national state, the political socialization and obligations of the citizen as an essential condition for civic incorporation and, facing the immanent danger of a fragmenting cultural pluralism, the pivotal role of a unitary national identity for the inclusion of immigrants. The key emphasis became *political* integration focusing on the relationship of the autonomous and moral individual citizen with the state seen as the universal representative of collective interest. This contrasted strongly, argues Favell, with a technocratic and piecemeal focus on socioeconomic *insertion,* which had dominated French integration policy until the mid-1980s. This previously dominant approach had responded to the presence of immigrants through a pragmatic but flexible concern for their basic welfare and social needs and by systematically playing down the issue as an object of political contest.

British policies of immigration and integration provide, maintains Favell, a perfect contrast to French republicanism. What has carried up the development of policies and elaborate institutional practices marked by cultural pluralism and antiracism in Britain since the mid-1970s is a general British philosophical tradition of pragmatism. It is profoundly alien in its self-understanding to the highbrow French intellectual mythology of *integration* concerned with the autonomous *citoyen* and its incorporation into a universalist political culture unfettered by particularist cultural allegiances.

A British mythology of peaceful evolutionary change based on piecemeal reform finds its supreme justification in T. H. Marshall's famous theory of citizenship, which has, in particular in post-Thatcher Britain, become the common ground for a broad consensus on integration and multiculturalism, running across the left-right spectrum of British politics. In contrast to France's focus on the nation and immigration control in terms of a defense of a unitary political culture and republican citizenship, argues Favell, the British conception of the nation is focused on political sovereignty and territoriality. The conception of citizenship as a top-down relationship between the sovereign and the subject has left a considerable space for cultural plurality in British society, which is not considered to interfere with politics or nationality. Potential principled political conflicts in terms of ethnicity and race are consistently manipulated off the political center by indirect rule-like policies of concession and control. British immigration control is justified in terms of pragmatic

concerns with the manageability of public order rather than in terms of a defense for a unitary public political culture and national unity.

Unlike many others, Favell does not settle in favor of one of these national philosophies and practices. Rather, they have both, he argues, and each in their way entered a stage of crisis marked by a growing disjunction between their respective national philosophy or creed (referring to Myrdal's *American Dilemma*) and social reality. In France, the dominant intellectualist concern with integration in ideal political terms has been instrumental in killing off earlier flexible forms of technocratic management of complex social issues and ethnocultural diversity. Its obsession with cultural unity of the nation has tended to unintendedly feed an exclusionary radical nationalism rather than furthering social integration. In Britain, on the other hand, a top-down focus on public order and pragmatic manipulation of ethno-cultural differences have tended to block actual political participation among ethnic minorities with consequences, which first became painfully clear with the Rushdie affair and the way it was conceived and acted upon by Britain's Muslim minorities. In both countries there is a growing gap between declared political ideals—in France republican *égalité*, in Britain antiracism and policies of racial equality—and a proliferating social exclusion with an ethnic and racial tinge. And both countries seem currently unable to break out of their national dilemmas, concludes Favell. He sees them as increasingly caught in the negative path dependence of their elaborate institutional frameworks and the broad national political consensus building that has previously been the basis of their relative success as well as the ground for their continued respective pretensions to serve as a model for others.

New policies of integration must go beyond national frameworks, concludes Favell. The ethnocentricity harbored in national philosophies of integration must be superseded in favor of the forging of elaborate transnational approaches. One powerful trans- or supranational framework is the European Union. But not only are both Britain and France unfit to serve as examples for common EU-policies on integration, concludes Favell. He sees their actual political behavior as belonging to the most serious obstacles for the development of any farsighted and coherent European policy. France and Britain are so obsessed by their own national philosophies that they belong to those political actors in the community that most consistently act to block any positive breakthrough for a common EU policy. They thereby act to exacerbate rather than to resolve an emerging wider European dilemma centered on ethnicity and the nation.

*Philosophies of Integration* is highly recommendable reading for anybody who wants to seek beyond established and sometimes locked, if not dogmatic, positions in the ongoing debate on immigration and citizenship in Western Europe and North America. The author is deeply acquainted with the national political systems and with the development of policies and institutions concerned with integration in both France and Britain. His argument is based on a critical over-time scrutiny of relevant documents and the political process. Favell joins the company of those who

like to see the escape from the deepening dilemmas, in which national policies of integration and citizenship are today obviously finding themselves, in political visions and regimes of a coming transnational order of citizenship. But he is sufficiently realistic as to not end up in any new dogmatism.

*Citizenship and Civil Society: A Framework of Rights and Obligations in Liberal, Traditional, and Social Democratic Regimes.* By Thomas Janoski. New York: Cambridge University Press, 1998. Pp. xii+316. $59.95 (cloth); $19.95 (paper).

J. M. Barbalet
*Australian National University*

The recent literature on citizenship is now extensive and continues to grow. Since the 1980s, there has been a consolidated interest in citizenship among social scientists of all stripes. This interest has been fed from clearly identifiable and diverse sources. The New Social Movements, for instance, influenced the practice of citizenship, and the way it is thought about, in various ways. Legal changes, especially concerning welfare entitlements and their conditions and relating to provisions available to migrant workers, especially in the European Union, also raise important questions about citizenship. This points to a problem: citizenship is a theme that yields to different concerns and treatments.

The book under review agrees with the above sentiments. Indeed, Janoski says that "citizenship is not just an area of interest, but, rather . . . a phenomenon that can be studied as a dependent variable and [also] used as an explanatory independent variable" (p. 228). This alone brings confusion and not just diversity to the literature on citizenship. The task Janoski has set for himself is to replace this lack of coherence with order.

*Citizenship and Civil Society* offers "a theoretical framework that is both coherent and capable of being operationalized" (p. 217). While much has been written about citizenship, Janoski complains that "little effort has focused systematically on social science theory" (p. 217). This book attempts to set things right. In attempting to do so, Janoski widely ranges over sociology, political science, and philosophy. But these different disciplines are lined up to point in the same direction: toward a testable comparative and historical theory of citizenship.

Drawing upon some distinctions in political theory, Janoski differentiates between liberal, traditional, and social democratic groups of countries. He also distinguishes between different types of rights (legal, political, social, and participation) and obligations. Against these are mapped characteristic patterns of rights and obligation and their historical development both in the short term (over decades) and in the long term (over centuries). What we have in this book, then, is a very broad cover of the central sociological problems of citizenship. The separate issues of

taxonomy of citizenship regimes, of characterization of rights and obligations, and of the historical and epochal development of citizenship, are each treated fulsomely and brought together within a framework that allows each part of the discussion to support the others.

There is much of value in this book that comes from its argument and evidence and also the presentation and organization of material drawn from the treatment of citizenship by others, especially in the historical chapters. The author notes throughout that the discussion is empirically grounded and testable. Indeed it is, in the main. But no serious discussion can operate exclusively on the level of data, of course. There are a couple of areas that this reviewer, at least, thought unconvincing in their conceptualization. As these are also matters the author tells us are important in his account, it is not inappropriate to mention them here. They relate to the conceptualization of rights and also obligations.

Janoski adds to T. H. Marshall's well-known checklist of rights—legal, political, and social—a fourth: participation rights. These latter "involve the state's creation of rights in private arenas" (p. 32). It is not clear, however, how these latter constitute a coherent set of rights at all. The need for them in Janoski's account may reflect a narrowness of definition of the three other types of rights.

A more serious problem, though, is that of obligation. Janoski notes that there has been reluctance on the part of most writers on citizenship to take obligations seriously; indeed, it is "the neglected aspect of citizenship" (see pp. 219–20). But one difficulty with the notion of obligation that Janoski does not consider is that it is simply not conterminous with the notion of rights. In an early chapter, it is suggested that to each right is a corresponding obligation, a position that cannot be sustained and is in fact dropped in a later chapter. But whereas citizenship rights are organizationally bounded, citizenship "obligations" are politically—even ideologically—bounded. This comes across quite starkly in the discussion of patriotism and especially "responsible patriotism." The account is highly normative. Even more pointed is the fact that noncitizens are equally obliged to fulfill the crucial obligations of taxation and social tolerance.

These critical remarks are not intended to damn *Citizenship and Civil Society* or discourage its being read. There is much more of value in this book than can be outlined here. In particular, the discussion of limited and generalized exchange is excellent, as is the treatment of the historical development of citizenship. This book is a valuable addition to the growing library of sociological accounts of citizenship.

*Formations of Class and Gender: Becoming Respectable.* By Beverley Skeggs. Thousand Oaks, Calif.: Sage, 1997. Pp. 192. $75.00 (cloth); $26.95 (paper).

Carolyn Steedman
*University of Warwick*

This book has already started to do work in the world, at least in the United Kingdom, where it was published 18 months ago. There has been no great fanfare of reviews; rather, by word of mouth, scholars in cultural, gender, and social studies have heard about Beverley Skeggs's exasperation at the disappearance of class from academic inquiry, recognized it as their own, and started to use her definitions and perspectives in their research.

Over a 12-year period Skeggs kept contact with a group of young women she first met when teaching on a community care course at a further education college in the northwest of England. She was supplementing her Ph.D. grant at first, and then the women became the focus of her study. She began the research "by just hanging around and talking to [them] as much as possible" (p. 22). Only later did the formal interviews and the tracking of families and life stories begin. The study expanded, so that this modernist ethnography, intent on understanding how subjectivities are constructed, is based on the experiences of 83 working-class women, all of them training for the helping professions. The ironies of employing the grand and abstract Foucauldian imperative are all the more forceful for being unstated here. The researcher studies the care of the self in women who were acquiring the "technologies" of a caring self, in order to spoon food into the mouths of the very young, the very old, and the very ill at one end, and to wipe bottoms and floors at the other, for very low wages indeed.

The theoretical perspectives that *are* stated, are the British cultural studies tradition, the kind of historical materialism that developed in the Birmingham Centre for Contemporary Cultural Studies under the influence of Pierre Bourdieu's sociology of class and embodiment. The most sustained argument is with feminist theory: class disappeared from feminist analysis in Britain sometime over the last 20 years; perhaps, suggests Skeggs, because feminist scholars just do not experience class in the way they do gender.

But class was "absolutely central" to the ways in which the young women lived their lives and experienced themselves, not least in their constant attempts to evade it: they had "a clear knowledge about their 'place' but were always trying to leave it" (p. 81). There is truly important evidence here of the extraordinary pains of knowing yourself never to be *quite right*: never having the right sort of body, the right sort of voice, the right sort of taste, of always being the wrong sort of person. Here in the United Kingdom these voices should be added to the fragmentary testimony we possess of the psychic costs of growing up on the other side

of the border, deposited by occasional visitors—like Henry Mayhew in the 1850s—to the country of the poor. In the United States this book could be most profitably read in conjunction with Richard Sennett and Jonathan Cobb's *Hidden Injuries of Class* (1977), which Skeggs cites, and also with Lillian Rubin's *Worlds of Pain: Life in the Working-Class Family* (1976), which she does not appear to know.

"Experience" is the key to Skeggs's definition of class, and the absence of E. P. Thompson's name from the discussion will strike historians reading this book as very odd indeed. So also might her version of "respectability." As an organizing concept that only emerged late in the research process it does a great deal of work in Skeggs's analysis. It is understood as "a central mechanism through which the concept class emerged," as an alternative formulation of possessive individualism, and sometimes as middle-class-ness (or non-working-class-ness) itself. It is a 300-year history of class and gender formation encapsulated: " 'Individuals' were the respectable, the moral, the worthy, the English, the White and the non-working class . . . a property of middle-class individuals defined against the masses" (p. 3). But the term probably cannot perform the conceptual or historical work that Skeggs wants from it. To say the least, it was an understanding of self in society that emerged in relation to men as well as to women, and when used with full historical resonance, really cannot be restricted to female experience, as Skeggs restricts it here. Moreover, respectability's power in historical analysis is that people knew what they were up to when they used the term to define themselves and others: 19th-century men and women knew if they were respectable, or if they were seen as respectable. On the evidence of these pages, it is not a word that Skeggs's young women used to describe themselves or anybody else. And Skeggs wants very badly to use concepts that have explanatory value for those "to whom they are meant to apply" (p. 166). There is a terrible irony here that is probably part of the story of exclusion she is trying to tell: finally, "respectability" is a framework imposed on her subjects' experience by the researcher.

The enduring value of this book will be its finely wrought account of how a number of young working-class women "do" subjectivity in late 20th-century England. When Skeggs says right at the end that "the women do not have coherent identities," and that they did not assume "that they had entitlements or that were even interesting" (p. 165) she is in fact telling us that all our conventional, modern theories of subjectivity were "designed for a different body in different circumstances," and that these theories continue to hide many working-class people from history. There is some very fine—and moving prose—at the end of a rather gracelessly written book, when Skeggs comes to know that she is writing tragedy, telling the great story of a world divided by ownership and exclusion. She is dealing with what the playwright David Hare once called "the immense engulfing sadness of capitalism" (*Asking Around* [Faber and Faber, 1993]).

*Camouflage Isn't Only for Combat: Gender, Sexuality, and Women in the Military.* By Melissa S. Herbert. New York: New York University Press, 1998. Pp. 205.

Rosemary Pringle
*Griffith University, Australia*

The 1970s was a period of radical change for the U.S. armed forces. The abolition of the draft and the adaptation to "peacetime" operations following the end of the Vietnam War encouraged the recruitment of women, not merely as auxiliaries but as fully integrated members. Experiments were made with sex-integrated basic training and, in 1976, women were finally admitted to the service academies to train as regular officers. Why then do they remain so marginalized? Melissa Herbert argues we must look beyond individual harassment or discrimination to the gender ideology that conflates masculinity with soldiering. Soldiering, she argues, is not only about going to war but about becoming "a man." Since most young men in the first three quarters of the 20th century had to complete some form of military service, basic training can be seen as the male equivalent of "finishing school" (p. 9).

The integration of women into the military poses more than the usual dilemmas faced by women in male-dominated occupations. How can a woman possibly be a soldier and a woman at the same time? How do women negotiate these contradictions? Following West and Zimmerman, Herbert approaches gender not as a given set of masculine or feminine attributes but as something we "do." She argues that military women have to strike a careful balance between femininity and masculinity, and that they develop quite conscious strategies about how best to represent themselves. In a situation where they may be judged too masculine simply by doing their jobs well, women may "do" femininity, for example, by using makeup or perfume, wearing earrings in uniform, putting on pumps instead of flat shoes, or donning skirts instead of pants. But they must be careful not to be "too feminine," which may be defined in relation to what are seen as the excesses of traditionally female behavior or, in this context, simply as a failure to adopt more traditionally male behaviours.

Herbert surveyed 285 women who have served in the military since 1976 and followed up with 14 in-depth interviews. Since sexuality interacts with gender in important ways, she has deliberately oversampled lesbian and bisexual women, predicting that their experiences will differ on the basis of their sexual orientation. While nearly half of her sample denied that they were under any great pressure to act as either feminine or masculine, a quarter thought there were pressures to act feminine and a third thought there were pressures to act masculine (surprisingly few identified pressures to do both). They also thought that women incurred penalties if they overacted and behaved in ways that were perceived as either "too" feminine or "too" masculine. Sexual orientation did not affect the likelihood of believing that there were penalties for being perceived

as too feminine but, not surprisingly, the lesbian and bisexual women were more sensitive about being penalized for being perceived as too masculine (p. 49). Regardless of sexual orientation, the main punishment for being too masculine was to be labeled a lesbian. Women seemed to be as likely as men to punish inappropriate gender behavior.

Herbert wants us to see women in the military not as victims but as strategists—yet their strategies seem limited and fraught with difficulties. The problem may lie as much with Herbert's social psychological approach as with the ways in which the women identify the issues. Despite her emphasis on gender as performance, she leaves little space to consider parody or diversity. It is a pity that she was not able to include a higher proportion of black women, whose ways of "doing gender" may be very different. The body is also strangely missing from the analysis, although how women acquire or transform the military "habitus" is a key question. The question of bodily difference also needs to be considered seriously in relation to the military customs that regulate the enforced intimacy of bodies in cramped and uncomfortable conditions.

While there is some acknowledgment of diverse military cultures, the overall emphasis is on their monolithic maleness. It would be interesting to know more about how different masculinities are formed both within and across the different armed forces and how this affects the strategies available to women. The author intimates that the air force ("brain over brawn") is more open to women and that the navy is the most difficult, but she does not go into any detail about how these different cultures and masculinities are produced. The account would have benefited from a comparison with other nontraditional occupations such as law, medicine, and the church, in which women have been able to "do gender" at a more institutional level. It is hard to believe that women would have had the success they have had in the military (and Herbert concedes they have made inroads that would have been unthinkable a decade ago) without some more effective strategies than the ones she discusses.

Herbert's book is rather repetitive, though short, and limited in its coverage. It would be useful to identify where the cracks in male dominance lie and how these can be widened. Repealing the law against lesbians and gay men is an important dimension, but only one, in the struggle to transform the military and women's place in it.

*Gender Vertigo: American Families in Transition.* By Barbara J. Risman. New Haven, Conn.: Yale University Press, 1998. Pp. xiii+189. $25.00.

Anita Ilta Garey
*University of New Hampshire*

Barbara Risman's book, *Gender Vertigo,* has two interrelated themes, one theoretical and one empirical. Her theoretical point is that "gender is itself a social structure" (p. 3); her empirical focus is on men and women

who, in the context of family life, do not follow traditional gendered scripts. Risman's research projects, the presentation of which forms the bulk of the book, are designed to show that gender is a structure rather than a set of internalized traits and to explore how the structure of gender operates in these nontraditional settings. The three empirical studies Risman presents are: her 1980–83 survey of single fathers, her early 1990s reanalysis of 1965–79 longitudinal data from the Career Development Study, and her more recent study of "fair families" that equitably share earning a living and household labor. One of her aims is to illustrate that people can move beyond gendered lives, through a period of "gender vertigo," and into a postgender society.

Risman's view of "gender as structure" draws primarily on the work of Judith Lorber and Robert Connell, and her understanding of the interactional production of gender draws on Candace West and Don Zimmerman's influential article, "Doing Gender" (*Gender and Society* 1 [1987]: 125–51) and on Connell's use of both West and Zimmerman and Anthony Giddens. Risman argues that, while she builds on these theories, she also moves beyond the theoretical work from which she draws by conceptualizing "gender as structure" that operates on three levels simultaneously: the individual level of gendered selves and gendered behavior, the interactional level of situational meaning and cultural accountability, and the institutional level of social organization and material constraints. Most feminist theorists of gender, including the ones mentioned above, do recognize that gender operates in this tripartite way, although each usually focuses on the explanatory value of one level in light of the limits of the others. Indeed, Risman herself calls for particular attention to the interactional and institutional domains, and the analysis of her research on single fathers and on women's employment decisions, for example, is presented as supporting the argument that structural dynamics at the institutional and situational levels are *more important* in determining behavior (such as "mothering" behavior by fathers or labor-force participation by women) than are sex-role socialization or gendered identities.

Risman is arguing against individual-level psychological approaches to the study of gender and arguing for structural approaches that include the interactional domain. Revising her early position that gender-neutral external structures are determinatively responsible for gendered behavior (p. 21), Risman articulates a structural theory that recognizes that individuals do have agency: "we must bring individuals back into structural theories" (p. 27). Few sociologists of gender would posit an actor-absent structuralist approach, and interactionist approaches treat interaction as being in a recursive relationship with structure. The theoretical message of the book thus appears to be aimed at those sociologists of gender who focus at the level of the individual and at those structural sociologists who do not yet see gender as a category of analysis.

In her research on "fair families," Risman analyzes gender as structure at all three levels of analysis: individual, interactional, and institutional. The 15 white, heterosexual-couple families that met the behavioral and

ideological criteria for being in the study (actually sharing equitably and *believing* that they both should and did share equally) were educationally and economically elite. While studies of a group selected for particular characteristics can illuminate social processes, they do not allow the comparisons that Risman sometimes makes between fair families and other families. The question, What allows these couples to construct fair marriages? (p. 126) is a comparative one that cannot be answered with a noncomparative sample, and her statement that "one factor that clearly differentiates these couples from most others is their joint commitment to equality and their challenge to the interactional gendered expectations usually built into heterosexual marriage" (p. 127) is not only comparative but also incorporates as a finding the original criteria for selection in the study.

Risman's main point in her discussion of the "fair families" findings, however, is that while gender structure shapes and constrains people's lives, people also actively shape and change the gender structure, and she uses Connell's conceptual framework (*Gender and Power* [Stanford University Press, 1987]) to examine gender in these families in the organization of work and production, authority and social control, and emotional connections. Risman treats each gendered domain in turn and concludes that "change must occur simultaneously at all levels of the gender structure," with women becoming economically independent at the structural level and men transforming their gendered selves to become nurturant and sensitive (p. 127). A particularly innovative component of the "fair families" study is Risman's inclusion of interviews with the children in these families.

Missing from the discussion of empirical findings is an analysis of the interaction between levels of the gender structure, an analysis that would be particularly well suited to the qualitative methods of observation and interviewing used in the fair families study. Indeed, the fair families study needs a book of its own in which to adequately present gender as structure at the interactional level as well as to analyze the interaction between gendered selves, gendered interaction, and gendered institutions.

*The Making of the Unborn Patient: A Social Anatomy of Fetal Surgery.* By Monica J. Casper. New Brunswick, N.J.: Rutgers University Press, 1998. Pp. xi+267. $50.00 (cloth); $20.00 (paper).

Charis Thompson
*University of Illinois at Urbana-Champaign*

*The Making of the Unborn Patient* is a fascinating book. Fetal surgery combines the social and scientific interest of pregnancy stories with the credulity-stretching and empathy-inducing characteristics of the frontiers of modern medicine. As Monica Casper discusses in the opening chapter, the topic has not received as much publicity as related areas of biomedi-

cine. This means that reading about fetal surgery comes as a shock in a way that such things as test-tube babies, postnatal surgery to separate conjoined twins, organ transplantation, prenatal screening, or gene therapy no longer do. That doctors cut open pregnant women's uteri, lift out the still-attached fetus, operate on the fetus, replace the fetus in the womb, and then replenish amniotic fluid and sew the womb and the woman's abdomen back up again, is still startling. And that some fetuses apparently heal from their surgery without scars (unlike their mothers) only serves to exaggerate the mystical allure of the specialty.

The book is peppered with phrases like "I interviewed a women who had undergone fetal surgery" (p. 26), where to understand the sentence at all, one must be able to make the switch from the woman to the fetus as the subject of the sentence. This elision of subjects contains the political import of Casper's study. In fetal surgery, the oft commented upon erasure of pregnant women achieved in the construction of fetal individuality is invidiously heightened by the fact that the pregnant women must themselves undergo major surgery for their fetuses to be patients. Casper delineates the connections between the ontology of the fetus as patient and right-to-life politics—made, for example, by the fetal surgery pioneer William Liley, who carried out the first fetal blood transfusion for Rh disease in 1963. She also documents the attempts by those involved two decades later in open fetal surgery to downplay the political and ethical implications of their work. Researchers use more and less routine procedures, such as experimenting with nonhuman animals, moving off mainland United States to Puerto Rico, and shifting their treatment between the bioethical categories of "experiment," "innovation," and "standard of care," for institutional review purposes. These tactics enable fetal surgeons to maximize access to patients, funds, and success, and to minimize their exposure to such potentially problematic bodies as ethics committees, feminist critics, antivivisectionist protesters, and cost-benefit minded insurers and hospital administrations. It is possible that the current shift away from open surgery to endoscopy will be accompanied by a new openness to public scrutiny.

Perhaps most interesting of all, Casper foregrounds the complex agency—complicit as well as resistant—of women implicated in fetal surgery. Combining the chronic problem of premature labor with the apparently atrocious results of fetal surgery compared to nonsurgical or postnatal interventions, the economic aspects of access to treatment and funding, the enforced postoperative disciplining of women patients' bodies to forestall premature labor, the link to secondary infertility and postnatal infant mortality, morbidity, and disability, and the relatively few indications for fetal surgery, the practice seems a prime example of being "from (the) mother's womb/Untimely ripp'd." Casper manages despite the termination of her formal access to patients to recount enthusiasm, ambivalence, and resistance among patients, social workers, and practitioners. (Why is there not a patient advocacy organization for fetal surgery through which she could have contacted patients independently, as

there is for many contested biomedical areas?) She uses compassion and respect in capturing the complex motivations and multiplicity of the actors involved. She also stands firm in her feminist demands to reframe fetal surgery as a women's health issue. I find this combination refreshingly honest and well accomplished.

There were a number of areas where I wanted more detail. The Lileys, Adamsons, Freda, and Karliss all got potted biographies, but what about Florence Frazer, one of the few women on the practitioner side of the story, and an important informant in the book? More too about the patients from whose moving written testimonials Casper excerpts would have been revealing as to the history of the specialty. Casper also gives short shrift to the institutional embedding of the specialty. She describes the conflicts between woman-oriented obstetricians and fetus-oriented fetal surgeons in the fetal treatment unit, but it is not until close to the end of the book that we learn the crucial and preemptive piece of information that fetal surgery is institutionally housed in pediatrics. She shows how fetal surgery can be assimilated to other examples of maternal sacrifice, but not how it is continuous with other nonmaternal ways in which people are called to sacrifice their bodies or parts thereof for others.

Casper argues that "fetuses are increasingly represented as free agents with their own needs and interests, and even their own doctors and lawyers." How general is the attribution of personhood to fetuses in and around fetal surgery? For example, are fetuses independently treated for *their* pain after the surgery? Fetuses in the contemporary imagination certainly do not achieve personhood or the right to life simply by being isolable, free-floating entities. In the War Remnants Museum's in Vietnam, for example, pickled free-floating fetuses are as strongly associated with men, war, teratogenesis and necrosis, as the gestating fetus is to their antonyms.

Casper's account shows that although biomedicine cannot determine whether or not the fetus is a person, yet it shows how it facilitates the attribution of personhood. The account also compellingly resists the simple maternal/fetal conflict paradigm that would pit those concerned with women against those concerned with the well-being of fetuses. *The Making of the Unborn Patient* is a highly valuable contribution to the story of the emergence of fetal personhood.

*Moral Panic: Changing Concepts of the Child Molester in Modern America.* By Philip Jenkins. New Haven, Conn.: Yale University Press, 1998. Pp. xii+302. $30.00.

Anthony Oberschall
*University of North Carolina at Chapel Hill*

A moral panic over a social problem occurs when official reaction to the problem is way out of proportion to the threat, fears are exaggerated and

wrongly directed, experts, professionals and advocates are unanimous, media reporting is sensational, and law enforcement catches a multitude of lesser deviance and crime in a tightened net of social control. Philip Jenkins argues that sex crimes against children in the United States have gone through several cycles of moral panic in the past 100 years.

The moral panics about children facing grave dangers of sexual abuse include changing conceptions of the child abusers or molesters, their motivations, their treatment, their progression from lesser to serious offenses, situations in which children are at risk, behaviors that are prosecuted, and their harmful consequences for victims. At one time or another, the child sex abuse issue got entangled with homosexuality, pedophilia, child pornography, ritual child abuse, incest, missing children, serial child murder, rape, family violence and physical abuse, and child prostitution. In Jenkins's view, sex abuse issues are socially constructed realities. Changing beliefs and values of the public, child advocates, professionals, and law enforcement, agents on child protection reflect shifts in values and ideologies associated with large social movements and demographic changes such as feminism, social reform, conservative moralism, the youth culture, and working mothers with children in day care. Social cause advocacy, the mass media, and the criminal justice system both feed on and fuel these trends. Sooner or later, highly visible child sex offenses trigger moral panics. The central factor enabling a moral panic is a coalition of conservatives and liberals normally opposed on public issues, which allows a rapid, compelling consensus to prevail.

In the Progressive Era, the coalition was of feminists allied to social workers and reformers to protect children, with law enforcement and corrections professionals eager to maintain social control on immigrants in cities, where much of the problem was located. With the end of immigration and the weakening of feminism after suffrage, the coalition came apart and child sex abuse concerns faded. From the mid-1930s, when sex psychopath statutes were passed and the medicalization of sexual offenses got a big boost, to the mid-1950s, the sex psychopath frame was dominant, only to be replaced in the 1960s by the child molester who was viewed as a much less dangerous offender than its predecessors.

The core of the book is the last two decades, where the evidence of panic is strongest. In the 1970s, feminists campaigning against rape and family violence, and the profamily conservative right, created a powerful coalition on child abuse, which soon encompassed sex abuse and branched into child pornography and associated issues. Major federal legislation, the Mondale Act, provided financing and an infrastructure to social work, therapy, and criminal justice professionals for reporting on, investigating, and treating child sex abuse. Mandatory reporting swelled statistics, and confidential reporting boosted unconfirmed allegations. Prosecution was made easier by redefining sex offenses, lowering standards of proof, and allowing children to testify. Television, newspapers, tabloids, cinema, and self-appointed experts saturated the public with child sex abuse. Child protection professionals and therapists participated

in conferences and seminars that changed the ideology and practice of child protection and the legal environment of prosecution. Practices producing false positives became accepted without challenge. Soon allegations and prosecutions spread to mass molestation in day care, ritual or satanic child sex abuse, pedophile and child pornography rings, and cyberporn.

Has the 1980s moral panic spent itself? Jenkins is not equivocal about it, and his demography-driven explanation for a value and ideology shift is unlikely. Instead, there may be feedback in the panic process: consensus and lack of challenge will create excesses and abuses, and the public, appeal courts, and academic researchers will react. Highly visible "miscarriages of justice" get publicized, and the public's sense of fairness and justice is offended. Thus a moral panic can be turned around through its internal dynamics, though it may take time. I believe it can be shown for the 1980s day care mass molestation panic.

Jenkins's book is very readable and cogently argued. It deals with a topic at the intersection of deviance, social change, social movements, culture, social problems, sex and gender, and perhaps others. It deserves to be widely read. One lesson for the reader: beware when there is a coalition of groups that usually are rivals who jointly create a consensus on a moral issue (or indeed any issue, such as war). Another lesson is that the panics were not grassroots driven, though in time ordinary people rallied to the cause. Jenkins is too ready to describe public attitudes from media coverage, but nowhere does he produce opinion surveys and polls. I suspect they would indicate less concern and "panic" than some advocates, therapists, and child protection and law enforcement professionals exhibit. It would be of interest to look at a negative case, for example, AIDS. Why was moral panic stopped early on? It could be that unlike the alleged child sex molesters, the AIDS sufferers are a highly organized and fighting group.

*For the Sake of the Children: The Social Organization of Responsibility in the Hospital and the Home.* By Carol A. Heimer and Lisa R. Staffen. Chicago: University of Chicago Press, 1998. Pp. xv+419. $50.00 (cloth); $18.00 (paper).

Monica J. Casper
*University of California, Santa Cruz*

We have heard a great deal recently about moral accountability as politicians exhort us to "take responsibility" for our lives. Republicans and Democrats alike celebrate individual responsibility and "bootstrap" mobility. The message is that people without jobs, money, or nuclear families must be irresponsible. What is missing from public debates is the kind of rich sociological understanding of responsibility offered by Carol Heimer and Lisa Staffen in this eloquent, important book.

Their basic argument is that responsibility is not simply an individual achievement rooted in psychology or morality, but rather it is socially produced within collective settings. They use the case of neonatal intensive care units (NICUs), where premature or impaired newborns are treated, to address questions about how responsibility forms. The NICU is a fascinating case study because its patients are especially needy. Those who take care of them, whether health care workers or parents, must be responsible if the babies are to get better and have decent lives. Good outcomes, as the authors argue, depend on coordinated care. Yet technical and moral difficulties abound as actors struggle over how best to take care of their fragile charges. In NICUs, boundaries are permeable and questions of who should take responsibility are ambiguous. Parents have an obvious claim on their newborns, yet they must exercise their agency on the high-tech turf of medical workers.

Heimer and Staffen posit five dimensions of being responsible: (1) taking other people's interests and needs seriously, (2) thinking about the long-term and planning for the future, (3) defining obligations diffusely, (4) using discretion (or flexibility) to meet contingencies, and (5) accepting the consequences of contingency and discretion. Drawing on data collected at two hospitals, the authors argue that while most parents take responsibility for their sick newborns, they are *encouraged* to do so by the NICU staff who employ social control mechanisms to ensure parental responsibility. For example, parents are labeled, rewarded, or punished depending on how accountable they are seen to be. Because neonatal care is hospital-based, parents are at an organizational disadvantage and are especially vulnerable to social control. At the same time, parents are able to act as agents on behalf of their newborns and engage in their own evaluations of medical staff. In the face of organizational barriers, parents can nonetheless shape the care given to their children by medical staff. Once parents bring newborns home, the boundaries shift again as parents become more legitimate while medical staff recede in importance. This entire process is regulated by the state, which has an interest in both adequate medical care and adequate parenting.

The book offers several findings of relevance beyond the NICU. First, organizational contexts are highly influential in creating the conditions under which people can act responsibly. People engage in moral behavior but not under conditions of their own choosing. The authors suggest that rather than asking when people should be accountable, we must examine instead what social arrangements encourage people to think of themselves as accountable. In addition, whoever is seen to "own" a particular problem is more likely to assume responsibility for it. In settings less ambiguous than the NICU, the lines of responsibility may be much more clearly drawn. Also, Heimer and Staffen suggest that social categories such as gender operate as cultural resources in many settings. Here, mothers were far more likely than fathers to participate in the care of sick newborns. From an organizational perspective, it makes sense to maximize available resources; because women are stereotyped as resources

by medical staff, they are encouraged more than men to parent responsibly. That the burden of newborn care falls predominantly on women should not surprise us. But if Heimer and Staffen are correct in asserting that responsibility is socially produced, changing the conditions under which it forms is a first step in changing who acts responsibly in the world.

Although insightful and engaging, the book is not without problems. Defining responsibility as "high quality compliance with norms" begs the question of how norms are established in the first place. Who gets to define what is responsible behavior in any given setting? While the authors lay out compelling criteria for what responsibility is, one wonders whether the parents in the study shared these criteria. What I missed was a discussion of what responsibility looked like from the parents' point of view, rather than what parents looked like from the authors' point of view. That Heimer and Staffen make assumptions about what responsibility is attests to the ubiquity of established cultural meanings available to politicians, sociologists, and others who want to talk about moral accountability. The book also contains a number of naive generalizations, such as "all parents love their own children," that threaten at times to undermine the analysis. This book should be of interest to sociologists who study families, gender, organizations, work, health care, ethics, and law, as well as to ethicists, policy analysts, social workers, and others.

# Family Structure, Educational Attainment, and Socioeconomic Success: Rethinking the "Pathology of Matriarchy"[1]

Timothy J. Biblarz
*University of Southern California*

Adrian E. Raftery
*University of Washington*

The effect of alternative family structures on children's educational and occupational success has been constant over the past 30 years. Higher rates of unemployment and lower-status occupational positions could account for the negative effect of single-mother families on children's attainment throughout the period. Children from single-father families and stepfamilies have consistently had lower attainments than children from both two-biological-parent and single-mother families. The influence of many other dimensions of children's family background declined from the 1960s to the 1980s but has declined no further since. Among six candidate theoretical frameworks, the findings are most consistent with an evolutionary view of parental investment.

The "pathology of matriarchy" hypothesis that came out of the Moynihan Report (1965) is that the absence of a father is destructive to children, particularly boys, because it means that children will lack the economic resources, role model, discipline, structure, and guidance that a father provides. Moynihan (1965) focused on the African-American family, but the publication came on the eve of what was to be 15 years (1965–80) of sustained, rapid increase in the divorce rate of non-Latino whites. Some researchers have carried the "pathology of matriarchy" view beyond the African-American family to the larger population. Popenoe (1996, p. 8),

echoing Moynihan, argues that "if we continue down the path of father-lessness, we are headed for social disaster."

Social science research has produced evidence both for and against the "pathology of matriarchy" view. Some studies using national samples show that children from single-mother families have lower attainments than children from two-biological-parent homes (Duncan and Duncan 1969; McLanahan and Sandefur 1994), while other studies, also using national samples, show that once other factors are taken into account, children from single-mother families do approximately as well as children from two-biological-parent families (Biblarz, Raftery, and Bucur 1997; McLanahan 1985). Some studies show that alternative family types—single-mother, single-father, and stepfamilies, for example—have similar, negative consequences for children (Dawson 1991); while other studies show that children from some kinds of nontraditional families have higher attainments, on average, than children from other kinds (Amato and Keith 1991b).

This article is about searching for order in this diversity of findings. We assess whether change over time in the effects of alternative families, and differences in researchers' decisions about which independent variables to include and leave out of models, can account for the discrepancies observed in the literature. We do this by tracking the relationship between alternative families and children's educational and occupational success over four decades—the 1960s, 1970s, 1980s, and 1990s—using four large nationally representative surveys. For each time period, we observe how the family structure/child outcome relationship changes depending on the other dimensions of respondents' family backgrounds that are taken into account.

We find that the effect of family structure on children's socioeconomic success has been constant over 30 years. In the 1960s and 1970s, the high rate of unemployment among single mothers could explain the negative effect of single-mother families on children's educational and occupational attainment. In the 1980s and 1990s, single mothers' low-status occupational positions—rather than their employment/unemployment status per se—could account for the negative effect of female headship. There were no significant differences by gender in the effects of alternative families—alternative families had the same effect on men and women. The influence of many other dimensions of children's family background declined from the 1960s to the 1980s, but this decline progressed no further between the 1980s and the 1990s.

The discrepancies found in the literature about the consequences of alternative families for children are due in part to different decisions about which variables are exogenous to the process and in part to different decisions about how to group alternative families. With or without various sets of controls, we find that over the past 30 years children from single-

father families, father/stepmother families, and mother/stepfather families have consistently had lower attainments than children from both two-biological-parent families and single-mother-headed families. Among six candidate theoretical frameworks, the findings are most consistent with an evolutionary view of parental investment.

## BACKGROUND

### Why Do Children from Alternative Families Have Lower Attainments?

Below we review the main theories of the effects of family structure on children and discuss the predictions made by each. These predictions are summarized in the second column of table 1.

*Sociological theory.*—Almost all existing theory about the consequences of family structure for children centers around the relationship between family type and resources. Under the general rubrics of "social structure and personality," or "social structure and psychological well-being," sociological theory—socialization, learning, and control theory—predicts that children from alternative families get fewer economic, social, and cultural resources, which help facilitate success.

TABLE 1

SUMMARY OF PREDICTIONS ABOUT FAMILY STRUCTURE EFFECTS ON CHILDREN'S
ATTAINMENT

| Theory | Static Prediction | Change Prediction |
|---|---|---|
| Sociological: | | |
|    Socialization-learning* ......... | 2BP > $\overline{2BP}$ | No change |
|    Control† ............................ .... . | 2BP > 2BP | No change |
| Economic‡ ....... .............. ......... | 2BP = STF = STM > SF > SM | No change |
| Evolutionary psychology§ ......... | 2BP > SM > STF | No change |
| | $\lor$ | |
| | SF > STM | |
| Selection bias (parental | | |
|    competence)‖ ....... ............. | 2BP > $\overline{2BP}$ | Increase in effect over time |
| Marital conflict• ....... . .. ...... .... | 2BP > $\overline{2BP}$ | Increase in effect over time |

NOTE —2BP = two-biological-parent family, $\overline{2BP}$ = any non–two-biological-parent family; STF = stepfather/biological-mother family; STM = stepmother/biological-father family; SF = single-father family; SM = single-mother family.

  * Baumrind (1980), Weinstein and Thornton (1989), Simons and Associates (1996).
  † Nock (1988), Astone and McLanahan (1991)
  ‡ Becker (1964, 1981)
  § Trivers (1972), Daly and Wilson (1995)
  ‖ Cherlin et al. (1991), Popenoe (1988, 1993)
  • Grych and Fincham (1990), Kline et al (1991)

Socialization theory emphasizes the essential role of parenting in shaping children's lives (Baumrind 1978, 1980; Parcel and Menaghan 1994). In the case of single-mother families, father-absence reduces the family's ability to provide optimal amounts of support and control to children (Astone and McLanahan 1991; Thomson, Hanson, and McLanahan 1994). The emotionally distressing event of a spouse's death or of divorce, coupled with responsibility overload, negatively impacts women's psychological well-being (Acock and Demo 1994; Crosby 1987; Simons and Associates 1996). This leads to inconsistent parenting (Hetherington, Cox, and Cox 1978), less supervision over children (Thomson, McLanahan, and Curtin 1992), parental authoritarianism (Bronfenbrenner 1979), and an expectation that children mature in ways inappropriate for their age (Weinstein and Thornton 1989). All of these undermine the healthy development of children.

Learning theory views the family as a primary site where children learn about how to get along in the society when they reach adulthood (Kohn 1969, 1983). Without a father, children will lack a male model of how to successfully achieve in market activity (Powell and Parcel 1997; McLanahan and Sandefur 1994). In two-biological-parent families, children learn about how authority relations are structured and how to successfully interact with authority figures (Nock 1988). This learning facilitates children's educational and occupational attainment. In father-absent homes, where mother/child relations run the risk of becoming more peerlike, and in stepfamilies, where stepparent/child relations may be defined more as friends than as parent/child, children will not learn these important skills.

At the same time that alternative families represent for children the removal of positive resources, they also covary with children's exposure to negative ones. Events that produce alternative family structures oftentimes involve a loss or trauma—either the death of a father (or less frequently, of a mother) or the divorce of parents followed by the father (most often) moving out of the home. These events are stressful in the short term and have the potential to impact children's long-term life course trajectories in negative ways (Amato and Booth 1997; Glenn and Kramer 1987; Mueller and Pope 1977; Wallerstein 1989). The number of changes in family configuration over the course of childhood is an important predictor of some outcomes, like the risk of having a child outside of marriage (Wu and Martinson 1993). For other outcomes, like children's educational and occupational attainment, exposure to alternative family structures (even from birth) seems to be more important than the number of disruptive family events experienced or durations spent in particular kinds of families (McLanahan and Sandefur 1994; Wojtkiewicz 1993). Children raised in single-parent families from birth, for example, have roughly the

same high school dropout rate as those who experienced a transition from a mother-father to a single-parent family following parental divorce.

Problems in child socialization and parental control will occur in a variety of types of alternative families that children may experience—single-parent families, stepfamilies, and so on. In the case of single-parent families, for example, the undermining of parental control is a structural consequence of the absence of the father from the residential home. Stepparents, on the other hand, have "only a limited license to parent" (Furstenberg and Cherlin 1991, p. 85), and the arrival of a stepparent into the residential home can create disruption and friction in intergenerational relations that can take years to resolve and adjust to (Cherlin and Furstenberg 1994). The main prediction of sociological perspectives is that the two-biological-parent family is generally the optimal form for the successful socialization of children in modern society and that children from any kind of alternative family will, on average, do less well.

*Economic theory.*—Economic theory proposes that socioeconomic success is partly a function of human capital (Becker 1964, 1981; Becker and Tomes 1986). Households act as singular units to maximize collective utility. Utility comes from commodities—like children—that are produced by investments in market activity and household services. The two-parent family is among the best-functioning forms in modern capitalist society because it allows for the provision of household services by one partner and economic resources (or market goods) by the other. This is a particularly efficient system for maximizing utility and, by extension, the human capital of children.

Since children's success depends on the economic resources and equivalent services that parents provide, children who spend most of their childhoods in a two-parent family (biological or stepfamily) will have the highest attainments because two parental figures are present to provide complementary resources. Single-parent families will yield less income from the market and have less time for the provision of household services. One parent cannot cover both market and nonmarket activities as successfully as two, and children from single-parent families, accordingly, will do less well. Among children from single-parent families, economic theory would predict that children from single-father families will do better than those from single-mother families because they will carry a substantial income advantage. Children who grow up in single-mother families will have the lowest attainments—mother-headed families average less than a third the income of two-parent families and about half the income of alternative father-headed families (Meyer and Garasky 1993; DaVanzo and Rahman 1993).

*Evolutionary psychology.*—The evolutionary perspective on the family

(e.g., Emlen 1997) gives more weight to the role of the mother than that of the father in determining children's fates, and it places special importance on biological relationships. The evolutionary view starts with the premise that mothers invest more of their resources in children than fathers. The survival (or, perhaps, in the modern context, the "well-being") of a given child is of greater interest to the mother than to the father, because more of the mother's than of the father's potential reproductive investment is tied up in any one child (Trivers 1972). Both parents attempt to balance investing in present children against investing in having additional children in ways that maximize their reproductive fitness. But because women's potential for having additional children is far lower than men's, they have a greater interest in making sure that the children they do have do well. Evolutionary psychologists sometimes depict motherhood as, in part, a strategic exercise in finding ways to secure material resources from sometimes reluctant fathers, whose reproductive calculus may be pulling them toward future children (and partners) more than present ones.

Like the others, this theory would predict that children from two-biological-parent families will have an advantage over those from other kinds of families. The father's average resource contribution to children will be less than the mother's, but not by much because humans have high male parental investment, and so children will benefit from the presence of the biological father. But in contrast to the economic model, for example, the evolutionary view predicts that children from alternative families will do better raised by a single mother than a single father. Children from single-mother families will also have advantages over those from stepfather/biological-mother families. The stepparent's concern with his own reproductive fitness is in competition with the stepchildren for the mother's resources, increasing the risk of abuse to children in families with a stepparent (Daly and Wilson 1996).

*Selection bias (parental competence).*—One of the unanswered questions in family structure research is whether the observed negative effect of alternative families on children represents a selection effect. One variant of the argument is that people who divorce, for example, are less stable or less competent at family life. Children who experience their parents' divorce do less well because their parents are less competent, not because of the divorce per se. Cherlin et al. (1991) found using longitudinal data that many child behavioral problems associated with divorce were actually present in the children prior to their parents' divorce. The divorce, like the negative child outcomes, may have been a consequence of some preexisting family dysfunction (but see Cherlin, Chase-Lansdale, and McRae 1998).

Parental competence involves role performance as parent in the family,

but it may also involve value orientations (Popenoe 1993, 1996). Modern parents, particularly those who divorce, may be less "child-centered" and hold weaker "family values" than those who do not divorce (Popenoe 1993). Alternative value systems may be another, related source of selection, causing a spurious relationship between family structure and children's attainment.

*Marital conflict.*—Another variant of the selection hypothesis is that the main detrimental effect on children is not divorce but family conflict. Divorce is often preceded by (and sometimes followed by) high levels of conflict. Marital conflict is hurtful to children. Children of divorce have lower attainments than children from two-parent families because they have had sustained exposure to their parents' discord (Amato and Booth 1997; Amato, Loomis, and Booth 1995; Glenn and Kramer 1987; Mueller and Pope 1977).

Both parental death and divorce create sadness, distress, and related problems for children in the short term. Negative effects on children's health, self-esteem, and school performance have been observed (Dawson 1991; Mauldon 1990; Kline, Johnston, and Tschann 1991; Mott, Kowaleski-Jones, and Menaghan 1997). In the long term, children from alternative families have lower average socioeconomic achievements—in education, occupation, and earnings—than those raised in two-biological-parent families (Duncan and Duncan 1969; McLanahan 1985; Amato and Keith 1991*b;* Biblarz and Raftery 1993; McLanahan and Sandefur 1994; Powell and Parcel 1997; Amato and Booth 1997). McLanahan and Sandefur (1994) present evidence of negative effects of alternative families on selected child outcomes across four national surveys.

Some scholars believe that this combination of reasonable theory and strong supporting evidence points to a scientific truth about the family—that the stable "intact" family remains the best-functioning form, at least for modern capitalist society. Glenn (1994) asserts that a plethora of evidence has led to "virtually unanimous" agreement among the best social scientists that alternative families are not in the best interests of children (but see Stacey 1996).

Disquiet about the New Consensus

Other findings and some theory give cause for uncertainty about the new consensus. In the classic 1973 Occupational Changes in a Generation (OCG II) survey, the negative impact of mother-headed households on sons' occupational attainment is small and entirely a function of women's disadvantaged employment and occupational positions (Biblarz et al. 1997). Powell and Parcel (1997) similarly find little adverse consequence of an alternative family structure for men's education, occupation, and

earnings in the Panel Study of Income Dynamics (PSID), although they do find adverse consequences for women. McLanahan's (1985) earlier analysis of the PSID showed that father absence had no significant effect on children's education once income is taken into account. Boggess (1998), also using the PSID, finds no effect of living with a single mother on children's likelihood of graduating from high school. McLanahan (1985, p. 898) concluded that her results "do not support the notion that the long term absence of a male role model itself is the major factor underlying family structure effects." In the National Education Longitudinal Survey, holding constant other factors, there are no differences between children from two-biological-parent homes and those from female-headed families in the odds of dropping out of high school or attending college (Painter 1998). Among the six family types included in Teachman, Paasch, and Carver (1997), "divorced mother" seems to be the only type to not directly increase children's odds of dropping out of high school, holding other factors constant.

## Why the Discrepancy in Findings?

Both the findings and conflicting nonfindings about the effects of family structure on children are based on large, national random samples, where results should be approximately the same. So why the discrepancies? One possibility is that the choice of exogenous "control" variables (e.g., race, gender, sibship size, parents' education, residence) as well as "intervening variables" (e.g., family head's income, employment status, occupational position, sibship size) varies across studies. Conclusions about family structure's effect may depend on the other variables that are taken into account.

For example, on average, the greater the number of siblings, the lower children's attainments (Blake 1989; Powell and Steelman 1993; Steelman and Powell 1989). If children from single-mother homes have fewer siblings than children from two-parent families, this would represent an advantage associated with the single-mother family structure. Studies that take away this advantage by controlling for number of siblings will show a stronger negative effect of single motherhood. Studies that do not control for siblings will show a weaker effect. Conversely, the higher the socioeconomic position of parents, the greater the socioeconomic attainment of children (Blau and Duncan 1967; Featherman and Hauser 1978). Single mothers have lower socioeconomic positions than the fathers (and some mothers) who head two-parent families. Studies that do not take parent's socioeconomic position into account will show a stronger effect of family structure.

A second possibility is that the effect of family structure has changed

TABLE 2

PERCENTAGE DISTRIBUTION OF CAUSES AMONG RESPONDENTS FROM ALTERNATIVE
FAMILIES

|  | Death of Parent | Divorce | Other |
|---|---|---|---|
| 1962 Occupational Changes in a Generation ........ .. .......... .. .. | 68 | 28 | 4 |
| 1973 Occupational Changes in a Generation .. .. .. .. .. ....... . | 60 | 36 | 4 |
| 1986–88 Surveys of Income and Program Participation . ...... | 42 | 53 | 5 |
| 1992–94 National Survey of Families and Households .......... | 33 | 62 | 5 |

NOTE.—Figures based on causes by birth cohorts derived from the General Social Surveys, 1973–96, pertains only to respondents from single-mother, single-father, mother/stepfather, and father/stepmother families.

over time (the change predictions of each theory are detailed below). Findings are not so much discrepant as simply pertinent to different periods. The evidence from national surveys spans at least 25 years, from about 1962 (e.g., Duncan 1967) to 1987–88, in the case of the first wave of the National Survey of Families and Households (NSFH). Among the cohorts represented over this time period, the main reason for families not to be two-biological-parent has shifted from the death of a parent to divorce (Bumpass and Sweet 1989).

Table 2 shows estimates of the percentage distribution of primary cause (death of parent, divorce, other) among adult respondents from alternative families in four national surveys (these surveys are analyzed later in the article). Because most of these surveys do not contain information on cause of alternative family structure, we calculated the cause distributions for each birth cohort in the 1973–96 General Social Surveys (GSS) and then applied the distributions to the samples listed in table 2, weighting by the number of respondents in each birth cohort (the oldest cohorts in the 1962 OCG survey were assigned the averages for the grouped "pre-1930 birth cohort" of the GSS).

Table 2 shows that, over the past 30 years, divorce has come to replace death of a parent among samples of adults from alternative family backgrounds. The replacement of cause across time is striking: In 1962, 68% of adults from alternative family backgrounds are estimated to have experienced the death of a parent and 28% experienced the divorce of their parents, compared with only 33% experiencing the death of a parent and 62% experiencing parental divorce in 1992–94.[2]

[2] The growth in the number of female-headed families between 1950 and 1980 has, for whites, been driven primarily by growth in the rate of marital disruption and, to a lesser extent, by growth in the rate of nonmarital childbearing (for African-Americans the pattern is reversed) (Wojtkiewicz, McLanahan, and Garfinkel 1990).

Some of the theories discussed above would predict that this shift in the cause structure accompanying cohort replacement should lead to an increase in the magnitude of the negative effect of alternative families over time (from the 1960s to the 1990s), while others would predict essentially no change. The change-over-time prediction of each theory is listed in the third column of table 1.

For the sociological theories, the keystone variable is family structure. Heightened risk of parental authoritarianism or neglect and inadequate models of authority relations are structural consequences of the absence of one parent, no matter what process gave rise to that absence. There should be few differences, for example, in outcomes between children from widowed-single-mother families and those from divorced-single-mother families because they share the same basic family structure. The sociological theories that emphasize structure accordingly would predict no change over time in the effect of single-parent and other kinds of alternative family structures, even as the process giving rise to the alternative structures shifts. Other sociological perspectives—like marital conflict—that focus more on family process than on family structure would make the opposite prediction (see below).

Economic and evolutionary theory also predict no change. From the evolutionary perspective, divorced and widowed single mothers have the same level of their own fitness tied up in the children, and so both types of mothers would have the same level of impetus to invest highly in their children. The presence of a nonbiological parent would negatively impact children, regardless of whether the biological father had died or the parents had divorced. The change in cause structure over time should not alter the implications for children of basic family forms.

Economic theory focuses on household structure and composition. It claims, for example, that two-parent households are more efficient at maximizing utility than one-parent households and so would also predict no change over time in the effect of alternative family structures.

These predictions of "no change" over time are based somewhat on the assumption "other things being equal." Some evidence, for example, while not conclusive, shows that widowed mothers and their children have greater access than divorced mothers to certain kinds of social supports

---

Never-married single mothers—while a fairly socioeconomically diverse group—are, on average, more likely than divorced single mothers to be young, have low levels of education, have children at younger ages, and participate in welfare programs (London 1996). Unfortunately, the GSS and the other data sets that we analyze do not allow us to distinguish among respondents from never-married and divorced single-mother homes. However, available evidence suggests that the socioeconomic attainments of children from both groups tend to be similar (McLanahan and Sandefur 1994).

and may enjoy advantaged socioeconomic positions (McLanahan, Garfinkel, and Ooms 1987; Acock and Keicolt 1989; Amato and Keith 1991a, 1991b; Holden and Smock 1991; Biblarz and Gottainer 1999; Sugarman 1993, 1995).[3] Economic theory would generally favor single-parent families with greater economic resources over those with fewer economic resources in predicting children's human capital and subsequent attainment. But, independent of resources, economic theory makes no prediction about difference in outcomes between groups from single-parent families produced by the death of a parent vis-à-vis divorce, nor does evolutionary theory (although for evolutionary theory it makes an important difference whether the single parent is the biological mother or father).

In contrast, the selection bias (or parental competence) perspective predicts that the magnitude of the negative effect of alternative families is growing over time, as the old cohorts who experienced a parent's death as children leave our samples, replaced by the new cohorts who experienced their parents' divorce. Widowed mothers would be more competent at family life than divorcees, on average, and so the children from alternative families in the earlier period should have done substantially better than those in the later period. Widowed mothers, who did not choose an alternative family structure for themselves and their children, will also have more traditional values and lifestyles than divorced parents, who did. These kinds of values will be positively functional for children's success in the society, more so than the alternative values of divorcees.

Marital conflict theory makes the same prediction—the negative effect of alternative families is growing over time. Children from alternative families in the earlier period would be expected to do substantially better than those in the later period because they will have been less likely, on average, to have had exposure to parental conflict. Children of divorce often have mixed and sometimes hostile feelings toward their fathers, for example, whereas children whose fathers died tend to develop a warm and positive inner construction of the deceased father (see, e.g., Rozondal's [1983] "halo" effect; see also Silverman, Nickman, and Worden 1992).

---

[3] Social Security survivor's benefit payments to single widowed mothers and their dependent children tend to be substantially higher than AFDC payments to other kinds of single mothers and, unlike AFDC, they carry no requirement of an assets test to collect benefits (Sugarman 1993, 1995). To the extent that widows with dependent children have, on average, more economic resources than their divorced and never-married counterparts, the replacement of widows with divorced and never-married mothers over time could produce a growing negative effect of the "single-mother" family type. Measures like the family head's employment status and occupational position—considered later in the article—will only partly take into account the differences in material resources available across these disparate family types.

If the predictions of the selection hypotheses (parental competence, marital conflict) are sustained, such a pattern of increase in the negative effect of alternative families over time would be particularly important in light of changes that have been occurring over time in the effects of other dimensions of children's family backgrounds. A transition from ascription to achievement, or a "march toward meritocracy," has been occurring in the United States, particularly since the 1960s. Featherman and Hauser (1978) find a modest decline over the 20th century in the strength of the relationship between the family's socioeconomic position and children's socioeconomic destinations. This decline became clearly evident between 1960 and 1970 and steepened from 1970 onward (Featherman and Hauser 1978; Hout 1984b, 1988; Grusky and DiPrete 1990; DiPrete and Grusky 1990). The effect of race on socioeconomic attainment has also declined substantially (Hout 1984a), particularly at higher education levels (Hout 1988). If the importance of family structure as a source of ascription has increased as that of the old sources has waned, the offset could produce a halt in the trend toward universalism. If the family's socioeconomic position is making a comeback as a source of ascription in the 1990s, as some recent evidence suggests (Hout 1997), this, coupled with a potentially growing family structure effect, could lead to a reversal of course—a trend toward growing inequality in the opportunity structure.

Interactions

Research on the consequences of family structure for children generally explores how alternative families directly affect children's health, psychological well-being, and socioeconomic success. Two potential consequences that have been subjected to less empirical testing involve interactions, or conditional relationships.

The first idea is that family disruption may reduce the intergenerational transmission of status. Drawing primarily on socialization and role modeling theory, Biblarz and Raftery (1993) argued that family disruption negatively impacts social-psychological dimensions of parent/child relations that facilitate family transmission. They found that the association between parent's occupation (socioeconomic status, occupational autonomy, and training, from Hout [1984b, 1988]) and son's occupation was weaker among sons from alternative families. This was extended in Biblarz et al. (1997), but both studies rely on early-1970s data (OCG II) that include only men. As far as we know, this hypothesis—that father absence may be more important as a disrupter of socioeconomic transmission than as a direct determiner of socioeconomic attainment—has not been tested further.

The second idea, raised by Duncan and Duncan (1969), is that children who grow up in alternative families may have a more difficult time capitalizing on their educational accomplishments or translating their education into occupational success. Duncan and Duncan (1969) note that the returns in occupational status to each additional year of education were greater among sons raised in two-biological-parent families than among sons from alternative families. Duncan and Duncan (1969) offer no speculation about the potential process involved, nor do they statistically test the observed differences in coefficients. To our knowledge, this possible consequence of family structure has not been explored further.

Fathers may be an important link to the public sphere for children. Getting a good job is partly a function of educational credentials, but it may also be a function of, among other things, parental connections and exposure to knowledge about how to secure jobs (McLanahan and Sandefur 1994). Assuming that parents generally have the greatest knowledge about the occupational stratum in which they are located, children from female-headed families would be at a significant disadvantage, because female heads occupy substantially lower occupational positions than male heads (Biblarz et al. 1997). With less exposure to the father-as-worker model, and potentially less access to his connections, children from alternative families may find it more difficult to get the payoff to education normally achieved by children from two-biological-parent families.

Approach to the Analysis

The first large-scale empirical attempt to assess the socioeconomic consequences for children of family disruption was Duncan and Duncan (1969; see also Duncan 1967). These articles have been cited over 100 times since their publication, including recently in Powell and Parcel (1997) and Hout (1997).

Duncan and Duncan (1969) found that men from female-headed families (as well as men from alternative male-headed families) had lower occupational achievements (measured as average scores on Duncan's 1961 socioeconomic index—SEI) than men from two-biological-parent families. Duncan and Duncan (1969, pp. 284–85) interpret the five-point gap in occupational achievement between sons from two-parent and those from female-headed families as possible support for the Moynihan view: "The analyses reported above lend some support to the notion that the son raised in a family headed by a female is handicapped with respect to occupational success. . . . The evidence in this paper obviously does not constitute 'proof' that the matriarchal family structure and the absence of a father are 'pathological.' For Negroes as for non-Negroes, however, the

indication that an intact family background facilitates occupational success is quite compelling."

In the analyses that follow, we go back to Duncan and Duncan (1969) and to the 1962 Occupational Changes in a Generation survey (OCG I) in an attempt to locate the sources, or mechanisms, by which female-headed families (and other types of alternative families) may have led to a reduction in children's occupational success.[4]

We then move forward in time, replicating the analyses on national surveys from each of the subsequent decades. Among the questions that guide our analyses are the following:

1. Do conclusions about family structure effects change substantially depending on the mix of control variables?
2. Do conclusions about family structure effects change substantially depending upon period?
3. Do the socioeconomic achievements of children from different types of alternative families vary substantially?
4. Which (if any) of the theoretical perspectives are supported in terms of both the static and change predictions (table 1) that they make?
5. Do alternative families exhibit a weaker level of intergenerational socioeconomic transmission than two-biological-parent families?
6. Do alternative families reduce children's ability to translate educational achievement into occupational success?

DATA

Data are the 1962 OCG I, the 1973 OCG II, the pooled 1986–88 Surveys of Income and Program Participation (SIPPs), and the 1992–94 second wave of the NSFH, matched to information contained in the first wave (1987–88).[5] The OCG I and OCG II surveys were mail supplements to

the March 1962 and March 1973 Current Population Surveys, respectively. They represent the civilian noninstitutional male population ages 20–64 in those years. Comprehensive discussions of the OCG surveys can be found in Blau and Duncan (1967) and Featherman and Hauser (1978).

The SIPPs are large longitudinal (9 waves of interviews over 36 months) household surveys administered by the U.S. Bureau of the Census, representing U.S. household members (both male and female) ages 15–64. The second waves of the 1986–88 SIPPs panels include a family background topical module that contains most of the variables needed for the present analysis.

The NSFH includes interviews with 13,008 respondents ages 19 and older who comprise a representative sample of the adult U.S. household population. In the second wave of the NSFH, 10,008 of the original respondents were reinterviewed five years later. We use the family background characteristics that respondents reported on at time 1 to predict their time 2 socioeconomic positions. Following Duncan and Duncan (1969), for all data sets we select U.S.-born respondents from nonfarm origins ages 25–64. Analyses based on the OCG surveys include only men. Analyses based on SIPPs and NSFH include both men and women.

The OCG surveys ask respondents who they lived with most of the time up to age 16. SIPPs asks respondents who they were living with at age 16. NSFH1 asks respondents who did not live with both biological parents from birth until age 19 who they lived with at each age, 0–19. For the NSFH, we use respondent's family arrangement at age 16.

Based on responses to these items, we constructed family types that were comparable across the four surveys: (1) two-biological-parent families; (2) alternative mother-headed families (comprised almost fully of single-mother families); (3) alternative father-headed families (including both single-father and father-stepmother families); and (4) mother-stepfather families. These "snapshot in time" indicators of a small set of primary family types generally do well at capturing the major effects of this dimension of family background (McLanahan and Sandefur 1994; Powell and Parcel 1997; Wojtkiewicz 1993).

The other independent variables include race, number of siblings, and the education status, employment status, and occupational status of the family head that respondents lived with during childhood.[6] Occupational

[6] To achieve comparability across surveys and to reflect the discrete nature of educational attainment, parent's education is classified as 0 = 0 years; 1 = 1–8 years; 2 = 9–11 years; 3 = 12 years; 4 = 13–15 years; 5 = 16 years; 6 = 17 or more years.

status was measured as Duncan's (1961) socioeconomic index (SEI) for the OCG surveys, and, for SIPPs and NSFH, Hauser and Warren's (1997) update of Duncan's SEI. The dependent variables are respondent's educational attainment (in years) and respondent's occupational status (SEI). The means (and percentages, where appropriate) of all the variables in the analyses, for each survey and family type, are shown in table 3.

RESULTS

Regression Models

Table 4 presents five models of occupational status (current SEI) for each of the four surveys. The BIC statistic is calculated to compare the relative fits of the models ($BIC = N \ln[1 - R^2] + p \ln[N]$, where $p$ is the number of independent variables and $N$ is the total sample size [Raftery 1986$a$, 1986$b$, 1995]). Following Logan (1996), one asterisk indicates effects for which there is positive evidence ($|t| \geq \sqrt{\ln N + 2}$), two asterisks indicate effects for which there is strong evidence ($|t| \geq \sqrt{\ln N + 6}$), and three asterisks indicate effects for which there is very strong evidence ($|t| \geq \sqrt{\ln N + 10}$), as defined in Raftery (1995). In a sample of 10,000 respondents, for example, positive evidence would require that the coefficient have $|t| \geq 3.35$, strong evidence would require that $|t| \geq 3.9$, and for very strong evidence, $|t| \geq 4.38$ (Raftery 1995).

The purpose of the analysis in table 4 is to observe the size and direction of change in the effect of family structure as other variables are added to the model, across the four surveys. We do this in order to assess how conclusions about the effects of family structure may depend on the other variables that are included or left out of models. Trends over time in the effects of independent variables are the focus of table 5.

The first section of table 4, based on the OCG I analyzed by Duncan and Duncan (1969), shows the effects (unstandardized betas from least-squares regression) of selected dimensions of family background on men's current occupational status. Model 1 (which is roughly equivalent to Duncan and Duncan [1969, table 2]) shows that men from alternative families hold occupations that are, on average, about 3–6 SEI points lower than those of men from two-biological-parent families.[7] In 1962, differences in

---

Respondents missing on family head's education or family head's occupation (e.g., SIPPs includes a category "family head had no paying job" when respondent was age 16) were assigned mean values specific to their race by family structure by gender subgroup, with dummy variables for "family head was not employed" and "family head's education not reported" included in the models.

[7] These coefficients differ somewhat from Duncan and Duncan (1969) because we separate alternative father-headed families and stepfather-headed families (Duncan and Duncan treat them as a single "male-headed" group) and because Duncan and

socioeconomic status between African-American men and others (mostly non-Latino whites) were huge (22 SEI points).

Models 2 and 3 take into account the family head's employment and occupational status, respectively. Model 2 shows that having a family head who was not employed in the paid labor force (or for whom no occupation was reported) reduces men's socioeconomic attainment (by 8 SEI points). Origin unemployment also explains most of the effect of single-mother families on son's SEI. There is no significant difference in the occupational attainment of sons from two-biological-parent families and those from single-mother families once the family head's employment status is taken into account. The effect of a single-mother family is further reduced by taking into account differences in family head's occupational position (model 3).[8]

Men from alternative father-headed families have a lower average SEI than men from two-biological-parent families, with or without taking family head's employment and occupational status into account. Stepfamilies are associated with lower socioeconomic attainment, but this effect is not statistically significant based on the BIC criterion.

Models 4 and 5 add family head's education and number of siblings, respectively, to model 3. As is well known, parent's education is associated with children's socioeconomic success, and the greater the number of siblings children have growing up, the lower their socioeconomic attainments (in this case, each additional sibling costs the respondent about 1.4 SEI points on average).

When family head's education and number of siblings are added to the model, the effect of an alternative mother-headed family increases (−0.59 to −1.41 to −2.36). The effect in the final model (−2.36) is not significant based on BIC but would be treated as significant by commonly used criteria ($t = -3.1$, $p < .01$; not shown in table). In the 1962 data, men from single-mother families had fewer siblings than men from other family types (see table 3). Their mothers also had slightly higher levels of education, on average, than other heads. In earlier models 1–3 (those without

---

Duncan treated all alternative female-headed families as a single group, whereas we include only alternative mother-headed families.

[8] Half of the men from mother-headed families in the OCG I reported no occupation for their family head, compared with only 5% of those from two-biological-parent families and 12% of those from alternative male-headed families (table 3). Blau and Duncan (1967) expressed concern that this puts researchers at particular risk of allowing origin unemployment to "carry" the effect of female-headed families in these data. The OCG II is well-suited to address Blau and Duncan's (1967) concern because the overlap between family structure and the family head's employment status is considerably less than in OCG I. Only 28% of OCG II respondents from single-mother-headed families report no occupation for their mothers.

TABLE 3

MEANS (UNLESS OTHERWISE NOTED) OF INDEPENDENT AND DEPENDENT VARIABLES BY CHILDHOOD FAMILY STRUCTURE AND SURVEY

| VARIABLE | Two-Biological-Parent | CHILDHOOD FAMILY STRUCTURE | | |
| --- | --- | --- | --- | --- |
| | | Alternative Mother-Headed | Alternative Father-Headed | Alternative Stepfather-Headed |
| **1962 OCG I.** | | | | |
| Occupational status (SEI) | 43 | 36 | 36 | 39 |
| Education (years) | 11.5 | 10.4 | 10.4 | 10.6 |
| African-American (%) | 6 | 17 | 13 | 11 |
| Socioeconomic origins: | | | | |
| Family head was not employed (%) | 5 | 50 | 8 | 17 |
| Family head's occupational SEI | 33 | 30 | 30 | 30 |
| Family head's education | 1.8 | 19 | 1.7 | 1.7 |
| Family head's education not reported (%) | 9 | 18 | 16 | 17 |
| Number of siblings | 3.8 | 3.4 | 4.1 | 3.5 |
| % | 83 | 10 | 3 | 3 |
| N | 9,750 | 1,204 | 409 | 335 |
| **1973 OCG II:** | | | | |
| Occupational status (SEI) | 46 | 41 | 40 | 38 |
| Education (years) | 12.5 | 11.5 | 11.1 | 11.3 |
| African-American (%) | 6 | 17 | 12 | 14 |
| Socioeconomic origins: | | | | |
| Family head was not employed (%) | 5 | 28 | 9 | 11 |
| Family head's occupational SEI | 35 | 27 | 28 | 32 |
| Family head's education | 2.0 | 1.9 | 1.7 | 2.0 |
| Family head's education not reported (%) | 5 | 11 | 9 | 12 |
| Number of siblings | 3.5 | 3.4 | 4.0 | 3.6 |
| % | 83 | 11 | 3 | 3 |
| N | 16,010 | 2,054 | 562 | 659 |

**1986–88 SIPPs:**

| | | | | |
|---|---|---|---|---|
| Occupational status (SEI) | 38 | 35 | 35 | 35 |
| Education (years) | 13.1 | 12.5 | 12.4 | 12.3 |
| African-American (%) | 8 | 22 | 11 | 16 |
| Female (%) | 51 | 54 | 49 | 56 |
| Socioeconomic origins: | | | | |
| Family head was not employed (%) | 7 | 36 | 12 | 13 |
| Family head's occupational SEI | 35 | 28 | 34 | 33 |
| Family head's education | 2.6 | 2.6 | 2.5 | 2.6 |
| Family head's education not reported (%) | 19 | 23 | 27 | 35 |
| Number of siblings | 3.1 | 3.2 | 3.5 | 3.3 |
| % | 78 | 11 | 4 | 8 |
| N | 27,017 | 3,815 | 1,220 | 2,776 |

**1992–94 NSFH2:**

| | | | | |
|---|---|---|---|---|
| Occupational status (SEI) | 40 | 36 | 36 | 38 |
| Education (years) | 13.4 | 12.7 | 12.6 | 12.6 |
| African-American (%) | 8 | 29 | 15 | 13 |
| Female (%) | 50 | 55 | 56 | 56 |
| Socioeconomic origins: | | | | |
| Family head was not employed (%) | 6 | 18 | 10 | 14 |
| Family head's occupational SEI | 36 | 29 | 36 | 34 |
| Family head's education | 2.9 | 2.9 | 2.8 | 2.9 |
| Family head's education not reported (%) | 8 | 12 | 14 | 23 |
| Number of siblings | 3.2 | 3.4 | 4.8 | 3.4 |
| % | 75 | 14 | 4 | 8 |
| N | 4,905 | 897 | 240 | 495 |

NOTE.—To achieve uniformity in measurement across surveys, family head's education is classified as 0 = 0 years, 1 = 1–8 years, 2 = 9–11 years; 3 = 12 years; 4 = 13–15 years, 5 = 16 years; 6 = 17 or more years OCG I and OCG II include men only. SIPPs and NSFH2 include men and women. For OCG I and OCG II, we use Duncan's (1961) SEI for 1960 census occupational titles. For SIPPs and NSFH, Hauser and Warren's (1997) update of Duncan's SEI for 1980 and 1990 census titles is used. "Family head was not employed" includes missing values on origin SEI, i.e., cases where no occupation was reported for the family head

TABLE 4

UNSTANDARDIZED BETAS FROM REGRESSION OF CURRENT OCCUPATIONAL STATUS ON DIMENSIONS OF FAMILY BACKGROUND: U.S.-BORN RESPONDENTS AGES 25–64, FROM NONFARM ORIGINS

| Independent Variable | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1962 OCG I ($N = 11,177$): | | | | | |
| Childhood family structure | | | | | |
| Two-biological-parent (reference) | ... | .... | ... | .. | ... |
| Alternative mother-headed | -4.76*** | -1.41 | -.59 | -1.41 | -2.36 |
| Alternative father-headed | -5.86*** | -5.71*** | -4.58* | -4.47* | -4.22* |
| Alternative stepfather-headed | -2.70 | -1.89 | -.54 | -.63 | -1.41 |
| Race (African-American = 1) | -22.47*** | -21.36*** | -15.64*** | -15.41*** | -14.85*** |
| Socioeconomic origins: | | | | | |
| Family head was not employed | | -7.71*** | -7.72*** | -5.92*** | -5.15*** |
| Family head's occupational SEI | | | .39*** | .30*** | .27*** |
| Family head's education | | | | 3.02*** | 2.59††† |
| Family head's education not reported | | | | -5.31*** | -4.67*** |
| Number of siblings | | | | | -1.39††† |
| Intercept | 44.28*** | 44.62*** | 31.12*** | 29.20*** | 36.01*** |
| $R^2$ | .0655 | .0725 | .1838 | .2051 | .2257 |
| BIC | -720 | -795 | -2,214 | -2,491 | -2,775 |
| 1973 OCG II ($N = 17,894$): | | | | | |
| Childhood family structure: | | | | | |
| Two-biological-parent (reference) | ... | ... | ... | ... | ... |
| Alternative mother-headed | -2.56** | -.94 | 1.20 | .84 | .17 |
| Alternative father-headed | -4.64** | -4.40** | -2.54 | -2.12 | -1.66 |
| Alternative stepfather-headed | -6.75*** | -6.28*** | -5.64*** | -5.49*** | -5.57*** |

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---|---|---|---|---|---|
| Race (African-American = 1) | −16.17*** | −16.05*** | −10.36*** | −10.37*** | −9.03*** |
| Socioeconomic origins: | | | | | |
| Family head was not employed | | −6.88*** | −6.80*** | −5.16*** | −4.04*** |
| Family head's occupational SEI | | | .36*** | .27*** | .24*** |
| Family head's education | | | | 2.60*** | 2.12*** |
| Family head's education not reported | | | | −7.86*** | −7.21*** |
| Number of siblings | | | | | −1.50*** |
| Intercept | 46.49*** | 46.79*** | 34.04*** | 32.17*** | 38.89*** |
| $R^2$ | .0325 | .0371 | .1348 | .1533 | .1744 |
| BIC | −552 | −628 | −2,532 | −2,899 | −3,341 |
| **1986–88 SIPPs (N = 27,143):** | | | | | |
| Childhood family structure, | | | | | |
| Two-biological-parent (reference) | . . . | . . . | . . . | . . . | . . . |
| Alternative mother-headed | −2.10*** | −1.44*** | .15 | −.94 | −.86 |
| Alternative father-headed | −2.48*** | −2.39*** | −2.28*** | −1.97** | −1.71* |
| Alternative stepfather-headed | −2.58*** | −2.44*** | −2.00*** | −1.69*** | −1.56*** |
| Race (African-American = 1) | −6.43*** | −6.33*** | −4.39*** | −3.66*** | −2.73*** |
| Gender (female = 1) | −1.94*** | −1.97*** | −2.09*** | −2.20*** | −2.21*** |
| African-American × female | 2.68*** | 2.68*** | 2.75*** | 2.56*** | 2.43** |
| Socioeconomic origins: | | | | | |
| Family head was not employed | | −2.47*** | −2.47*** | −.22 | −.93 |
| Family head's occupational SEI | | | .29*** | .18*** | .17*** |
| Family head's education | | | | 1.51*** | 1.35*** |
| Family head's education not reported | | | | −3.78*** | −3.72*** |
| Number of siblings | | | | | −.70*** |
| Intercept | 38.89*** | 39.06*** | 28.70*** | 27.52*** | 30.45*** |
| $R^2$ | .0215 | .0240 | .0871 | .1068 | .1196 |
| BIC | −529 | −588 | −2,392 | −2,964 | −3,345 |

## TABLE 4 (Continued)

| Independent Variable | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **1992–94 NSFH2 ($N$ = 4,757):** | | | | | |
| **Childhood family structure.** | | | | | |
| Two-biological-parent (reference) | ... | ... | ... | . | ... |
| Alternative mother-headed | -2.90*** | -2.59*** | -1.10 | -2.07 | -2.21* |
| Alternative father-headed | -4.18* | -4.10* | -4.36** | -4.08* | -3.29* |
| Alternative stepfather-headed | -2.23 | -1.95 | -1.44 | -1.00 | -.92 |
| Race (African-American = 1) | -7.63*** | -7.41*** | -5.49*** | -3.89** | -3.16* |
| Gender (female = 1) | -1.75** | -1.74** | -1.87*** | -1.66** | -1.66** |
| African-American × female | 3.57 | 3.59 | 3.83 | 3.56 | 3.80 |
| **Socioeconomic origins:** | | | | | |
| Family head was not employed | | -3.56*** | -3.57*** | -2.41 | -2.17 |
| Family head's occupational SEI | | | .28**** | .16**** | .16*** |
| Family head's education | | | | 1.50**** | 1.37**** |
| Family head's education not reported | | | | -5.13**** | -4.72**** |
| Number of siblings | | | | | -.57**** |
| Intercept | 41.36*** | 41.53*** | 31.14*** | 28.58*** | 31.00*** |
| $R^2$ | .0294 | .0334 | .0939 | .1268 | .1353 |
| BIC | -91 | -102 | -401 | -560 | -598 |

* $|t| \geq \sqrt{\ln N} + 2$
** $|t| \geq \sqrt{\ln N} + 6$
*** $|t| \geq \sqrt{\ln N} + 10$.

controls for parent's education and number of siblings), these advantages of single-mother families were serving to offset other kinds of disadvantages. Taking away these advantages through statistical control in models 4 and 5 increases the negative impact of mother headship. Alternative father-headed families are associated with lower socioeconomic attainment regardless of control variables, while the effect of stepfamilies remains not significant.

The second section of table 4 goes through the same steps as above, on the 1973 OCG II. As in the OCG I, sons from alternative mother-headed families had lower occupational attainments than sons from two-biological-parent families mainly because they were more likely to have a family head who was not employed. This is reflected in the change in the effect of a mother-headed family from −2.56 in model 1 to −0.94 in model 2. It is interesting that, in both the OCG I and OCG II, among men who report that their family head had *any* paying job (regardless of the type of job), men from single-mother families held socioeconomic positions that were not significantly different from those held by men from two-biological-parent households. Taking into account parent's education and number of siblings does not, in the case of the OCG II, lead to a reemergence of a negative effect of single-mother families.

The occupational attainments of sons from alternative male-headed families are significantly lower than those of sons from two-biological-parent families. In the case of stepfamilies, the negative effect is strong regardless of control variables. The effect of alternative father-headed families is weaker and drops below BIC-based significance in model 3 onward.

The third and fourth sections of table 4 present the same analysis using the 1986–88 SIPPs and the 1992–94 NSFH2. These surveys include both men and women. We added gender (female $= 1$) to the models and tested all possible interactions involving gender and the other variables. There were no significant differences by gender in the effects of alternative families—alternative families had the same effect on men and women.

Gender is important in two ways. First, gender has a "direct" effect on occupational attainment: the average occupational status of women is lower than that of men. Second, the effect of race on occupational attainment depends on gender. Independent of other factors (model 5), African-American men average 2.73 SEI points lower than other race groups in SIPPs and 3.16 SEI points lower in the NSFH2. Among women, race has no significant direct effect on occupational attainment ($-2.73 + 2.43 = -.30$ in SIPPs and $-3.16 + 3.80 = .64$ in NSFH2).

The effects of alternative families roughly parallel those observed in the OCG I. Taking into account the family head's employment and occupational status reduces the effect of single-mother families to insignifi-

cance in both surveys. Relative to the OCG surveys, in both SIPPs and the NSFH the negative effect of single-mother families on children's socioeconomic attainment is more a function of mothers' disadvantaged occupational position and less a function of their disadvantaged employment status. In the OCG surveys, differences in employment explain most of the single-mother family effect. In SIPPs and NSFH, differences in occupational status explain most of the effect. Adding family head's education and number of siblings leads to the reemergence of a significant negative effect of single motherhood in the NSFH2. In SIPPs, the effect ($-.86$; $t = -3.0$; $p < .01$) is not significant based on BIC. Coming from any alternative male-headed families is generally associated with lower socioeconomic attainment in both of these surveys, with or without various sets of controls.

By including in table 4 only respondents who are presently working in the paid labor force (and hence have a current occupational status), we may be selecting the more successful children from alternative families. This would not be an issue in examining educational attainment, since virtually no one in any of the surveys is "missing" on education. Also, education is generally an earlier life course event than current occupation, and it tends to be completed in young adulthood. We might expect childhood family structure to have stronger effects on events that take place earlier in the life course.

The appendix replicates the analysis in table 4, treating education rather than occupation as the dependent variable. The main patterns are the same as those observed in table 4, only stronger and more clearly defined: family head's employment status and occupational status fully explain the negative effect of single-mother families in all four surveys. Controlling for family head's education and number of siblings brings the effect of single-mother families back, though below its original level. Stepfamilies and alternative father-headed families have negative effects on children's attainment, regardless of control variables. The negative effects of alternative father- or stepfather-headed families are generally larger than those of alternative mother-headed families.

Trends over Time

The analysis in table 5 focuses on trends in the effects of dimensions of family background over time. Our main interest is determining whether family structure effects depend on period. We merged the four surveys (selecting men only, since they are represented in all four time periods), created a year/period variable comprised of three dummies representing 1973, 1987, and 1993 (with 1962 as the referent), and then created interaction terms between the year dummies and each of the independent vari-

ables. Based on our inspection of the time-variable interaction coefficients, and theoretical models of change in patterns of inequality in the United States, we constrained the time changes in the effects of each independent variable to follow one of four forms: (1) no change over time (1962–93) in the effect of the independent variable on son's socioeconomic attainment; (2) steady (yearly) increase over time in the effect of the independent variable; (3) steady decline over time in the effect of the independent variable (called the "march-to-meritocracy" model); and (4) steady decline in the effect of the independent variable from 1962 through 1987, and no decline since (called the "speed bumps" model, from Hout [1997]).[9]

Table 5 presents two models. The "full model" includes the main effect of each independent variable plus all interaction terms involving year/period. The interaction coefficients are the differences in the size of the effects of the independent variable between 1962 and each subsequent period. For clarity of presentation, we show the actual effects implied for each year, rather than the difference scores. For 1962, $t$-values shown are for the test of whether the effect of the variable was significant in that year. For all other years, $t$-values shown are for the test of whether the effect of the variable is significantly different from its effect in 1962.

For example, the panel involving race shows that, independent of other factors, in 1962 African-American men held occupations that were, on average, about 15 points lower on the SEI than men of other races. In 1973, they held occupations that were, on average, 9 points lower. In 1987, they were 2.7 points lower, and in 1993, 3.4 points lower. The $t$-value −18.1 shows that race had a highly significant effect on occupational attainment in 1962. The $t$-values 5.7, 12.3, and 7.7 show that the effect of race in each subsequent period is significantly lower than its effect in 1962. The BIC-best model then finds the best way to describe change in the effects over time, from the candidate patterns discussed above.[10]

The effects of alternative families on children's attainment have remained constant over 30 years (the best model is that of "no change in

---

[9] The years of our survey data are 1962, 1973, 1986–88 (1987 midpoint), and 1992–94 (1993 midpoint). To estimate the various types of trend models, we created period scales and then multiplied them by independent variables. For the march-toward-meritocracy model, the period scale was 1962 = 0, 1973 = 11, 1987 = 25, and 1993 = 31. This form is used to test predictions that the effect of the independent variable(s) declined (or increased) steadily over time (i.e., by the same amount each year between 1962 and 1993). For the "speed bumps" model, the scale was 1962 = 0, 1973 = 11, 1987 = 25, and 1993 = 25. This constrains the effect of the independent variable to be constant from 1987 onward.

[10] The model is "BIC-best" within the constraint that some form of each of the independent variables be included in the model. We do this to provide maximum information on effects both significant and not.

345

TABLE 5

TRENDS OVER TIME IN THE EFFECTS OF DIMENSIONS OF FAMILY BACKGROUND ON
CURRENT OCCUPATIONAL STATUS: U.S.-BORN RESPONDENTS AGES 25–64 FROM
NONFARM ORIGINS

| Independent Variable | Full Model | t-value | BIC-Best Model | t-value |
|---|---|---|---|---|
| Intercept .............. .......................... ............. | 36 01*** | 60.7 | 36.00*** | 81.1 |
| Childhood family structure: | | | | |
| Two-biological-parent family | | | | |
| (reference) ...... ............... ... ...... . | . . | . . . | . . . | . . |
| Alternative mother-headed family | | | | |
| Effect of variable in: | | | | |
| 1962 ..... ....... . ...................... ............. | −2.36 | −3.1 | | |
| 1973 . ....... ........................................ | .17 | 2.8 | | |
| 1987 . . .... ................... .. .................... | −.84 | 1.7 | | |
| 1993 ...................... .. ...................... | −1.95 | .3 | | |
| Best model of change: | | | | |
| No change in effect over time ..... | | | − 87 | −2.8 |
| Alternative father-headed family: | | | | |
| Effect of variable in· | | | | |
| 1962 ................. ........................ . | −4.22* | −3.7 | | |
| 1973 ...................... .................... ....... | −1.66 | 1.7 | | |
| 1987 ........... ....................... .. . ..... | −1.61 | 1.9 | | |
| 1993 ....... .............................. ............ | −2.45 | .8 | | |
| Best model of change: | | | | |
| No change in effect over time ..... | | | −2.20** | −4.4 |
| Alternative stepfather-headed family: | | | | |
| Effect of variable in: | | | | |
| 1962 ............... . ................... . ......... | −1.41 | −1.1 | | |
| 1973 ...... .. ........................... ............... | −5 57 | −2 8 | | |
| 1987 .................... .. .......................... | −1.36 | 0 | | |
| 1993 .... .. ..... ..... .......................... ... | −2.18 | −.4 | | |
| Best model of change: | | | | |
| No change in effect over time ... . | | | −2.44*** | −5.8 |
| Race (African-American = 1): | | | | |
| Effect of variable in: | | | | |
| 1962 .................................... ............. | −14.85*** | −18.1 | −14 66*** | −22.1 |
| 1973 .. ............... .. .. .. .................. . | −9.03*** | 5.7 | | |
| 1987 ....... .. ................................ ........ | −2.65*** | 12.3 | | |
| 1993 ...................... ...................... . | −3.41*** | 7.7 | | |
| Best model of change: | | | | |
| Speed bump model ..................... | | | .48*** | 13.4 |
| Family head's occupational SEI | | | | |
| Effect of variable in: | | | | |
| 1962 . ... ... ......................... ............... | .27*** | 24.4 | .28*** | 30.0 |
| 1973 ................................................ | .24 | −2.1 | | |
| 1987 ... ................... .................... ..... | .20** | −4.1 | | |
| 1993 .. ......................... ................ . | .18 | −2.6 | | |

TABLE 5 *(Continued)*

| Independent Variable | Full Model | *t*-value | BIC-Best Model | *t*-value |
|---|---|---|---|---|
| Best model of change: | | | | |
| March-to-meritocracy ............ ... . | | | −.003*** | −4.6 |
| Family head was not employed: | | | | |
| Effect of variable in: | | | | |
| 1962 ........................... . ....... ............ | −5 15*** | −6.6 | −6.19*** | −10.4 |
| 1973 .............. ...... ....................... ... | −4.04 | 1.1 | | |
| 1987 ........ ... .. ..................... ...... . .. | −.77*** | 4 6 | | |
| 1993 ... ............................. .. .. ........ . | −2.00 | 2.0 | | |
| Best model of change: | | | | |
| Speed bump model ...................... | | | .224*** | 7.2 |
| Family head's education: | | | | |
| Effect of variable in: | | | | |
| 1962 ...... .. ........................... ..... .. .. | 2.59*** | 13.8 | 2 64*** | 17.6 |
| 1973 ... ............................. ..... ......... .. | 2.12 | −2.1 | | |
| 1987 .. ........................ ...... ................ | 1.24*** | −5.8 | | |
| 1993 ........................... .......... ............. | 1.56 | −3.0 | | |
| Best model of change. | | | | |
| Speed bump model ................... . . | | | −.053*** | −6 4 |
| Family head's education not reported: | | | | |
| Effect of variable in: | | | | |
| 1962 ..................... ... ... .. ................. | −4.67*** | −6.8 | | |
| 1973 ........ .. .. ................ ......... ...... . | −7.21 | −2.7 | | |
| 1987 .. . .. .......................... ......... .. .. . | −3.39 | 1.6 | | |
| 1993 ....... ... .................. ...... .. ......... | −5.70 | −.9 | | |
| Best model of change: | | | | |
| No change in effect over time ..... | | | −4.39*** | −15.3 |
| Number of siblings: | | | | |
| Effect of variable in: | | | | |
| 1962 ... ... .............................. .. .. | −1.39*** | −17 3 | −1.47*** | −30.7 |
| 1973 ..... ....................... ....... .. .. . ... | −1.50 | −1.1 | | |
| 1987 ................. .. .. ......................... | − 72**⁻ | 6.6 | | |
| 1993 .............. .. ...... ......................... | −.47*** | 5.6 | | |
| Best model of change: | | | | |
| March-to-meritocracy since 1973 | | | .053*** | 10 8 |
| $R^2$ ... .. ........................ ... .. .. .. .. ........... | 1939 | | .1926 | |
| *df* ............................ ...... .. .............................. .. . | 39 | | 17 | |
| BIC ..... . .. ................... ....... .. ...... . ........ | −9,455 | | −9,617 | |

NOTE —For 1962, *t*-value shown ıs for test of whether the coefficient ıs sıgnıficant in that year. For all other years, *t*-value shown ıs for test of whether the effect observed ın that year is sıgnıficantly different than that observed in 1962. All models include dummy variables for period For march-to-meritocracy and speed bump models, the coefficients shown are the change per year

    * $|t| \geq \sqrt{\ln N + 2}$.

    ** $|t| \geq \sqrt{\ln N + 6}$.

    *** $|t| \geq \sqrt{\ln N + 10}$

effect over time"). There is fluctuation in the family structure coefficients across years (of the sort we would expect from sampling error), but there is no clear pattern. Further, the *t*-values show no significant differences in the coefficients between years. BIC indicates that the fluctuation across periods should be treated as noise rather than as something more systematic.[11]

Hence, over 30 years of time, all else being equal (including number of siblings, and so on), a single-mother family has no effect on children's occupational attainment. The −0.87 SEI point ($t = -2.8$) associated with the single-mother form is not statistically significant by BIC standards (where in a sample of over 45,000 respondents, the minimum $|t|$ for indication of any positive evidence of an effect would be approximately 3.6). On the other hand, there is strong evidence that, independent of other factors, coming out of single-father families and stepfamilies is associated with lower occupational attainment. It is interesting that the effects of the alternative male-headed families are significantly more detrimental than the effect of single-mother families.[12]

The ordering of the values of the alternative family structure coefficients across these data sets suggests that two-biological-parent families have greater value than alternative mother-headed families, which, in turn, have greater value than alternative father- or stepfather-headed family. This was approximately the ordering observed in Biblarz et al. (1997) that led to their construction of a "distance from mother" scale, and it is also the ordering predicted by the evolutionary psychology perspective. To test whether this ordering is supported by the data, we refit the BIC-best model displayed in table 5 but treated family structure as a scale equal to 0 for two-biological-parent background, 1 for alternative mother-headed family, and 2 for alternative father- or stepfather-headed family. This model (not shown in table 5) explained the same proportion of variance (0.1926) as that explained by the BIC-best model shown in table 5, but it yielded two additional degrees of freedom and hence was more parsimonious (15 predictor variables as against 17). Accordingly, this turns out to be the strongest model given the data (BIC = −9638, as against −9617 for the BIC-best model shown in table 5).[13]

Other dimensions of family background decline in their ability to deter-

---

[11] We also conducted a trend analysis of the "total" effects (model 1 of table 4) of the family types, and the results were the same: the best model says that the effects of alternative family structures have remained constant over the time period (table available from authors).

[12] The *t*-values are 2.24 and 2.99 for difference between −0 87 and −2.20, and between −0.87 and −2.44, respectively.

[13] This ordering also fits the data best with education as the dependent variable.

mine children's destinations but then hit discernable "speed bumps" (Hout 1997) in the late 1980s. From 1962 through 1987, the effects of origin education, origin unemployment, and race declined steadily. Then, from 1987 to the present, the decline stopped. For example, the BIC-best model of the race effect, the "speed bumps" form, says that in 1962, African-American men held occupations that were, on average, 14.66 points lower on the SEI than men of other races. Since that time, the race gap diminished by 0.48 points every year through 1987 (hitting a low of −2.66, or −14.66 + [.48 × 25 years]), at which point it hit a "speed bump," and no further decline in the significance of race has been observed. In fact, in the "raw" data the effect of race on occupational status actually increased from 1987 to 1993 (from −2.65 to −3.41). This empirical observation fits with wide speculation and other kinds of evidence that there has been a recent rise in race inequality (Moss and Tilly 1996).

The effect of origin unemployment follows a similar pattern. Having a family head who was not employed mattered more for children's own achievement in 1962 than it did in 1987. But then, from 1987 to the present, this "march toward meritocracy" stopped (and even reversed in the raw data). This trend is particularly important because a larger proportion of new cohorts will come from alternative families, and children from alternative families have substantially greater odds of having a family head who is not employed.

Other dimensions of family background exhibit a "march toward meritocracy" that is as yet uninterrupted. The number of siblings that children have growing up matters less now than before. The effect of family head's occupational status has also declined from 1962 to the present. Independent of other factors, in 1962 each point increase in the family head's SEI produced a 0.27 point increase in the son's occupational status. By 1993, that effect had diminished to 0.18.

It was on this variable that Hout's (1997) "speed bumps" analogy was based. Hout finds suggestive evidence that the decline in the effect of origin SEI hit a "speed bump" in the late 1980s. While on face our results do not support this (note the way that the direct effect of origin SEI declines over time from 0.27 to 0.24 to 0.20 to 0.18 in our data), imposing a "speed bumps" form on the pattern of change in the origin SEI effect observed in our data leads to a fit based on BIC (not shown in table 5) that is identical to that shown in table 5. That we find the "speed bump" model supported for another central dimension of socioeconomic origins—family head's education, coupled with the "speed bumps" observed in the effect of race and origin unemployment—suggests the possibility of a trend in the United States toward (re)emerging inequality in the opportunity structure going into the 21st century.

Interactions

In the final steps of the analysis, we add to the BIC-best model (shown in table 5) terms representing interactions between each family type and origin SEI to test the hypothesis that alternative families exhibit a weaker level of intergenerational occupational transmission. All the interaction terms were statistically significant at the 0.01 level. There were no significant differences in the interaction terms over periods.

Coefficients of this type are difficult to interpret by simple inspection. In figure 1, we show the value of son's SEI predicted by the interaction model for three levels of family head's SEI across the four family types. All of the other variables in the model (siblings, period, origin unemployment, and so on) have been adjusted by multiplying their coefficients by their mean values and summing.

An inspection of figure 1 shows that the association between family head's SEI and sons' SEI is strongest among two-parent families and weaker among alternative families (weakest of all in stepfamilies). Put differently, parent's SEI plays a more important role in determining son's SEI in a two-parent family context than in other kinds of families. This is reflected in differences in the steepness of the slopes shown in figure 1.

In two-parent homes, moving from the low (10) to the high (70) end of
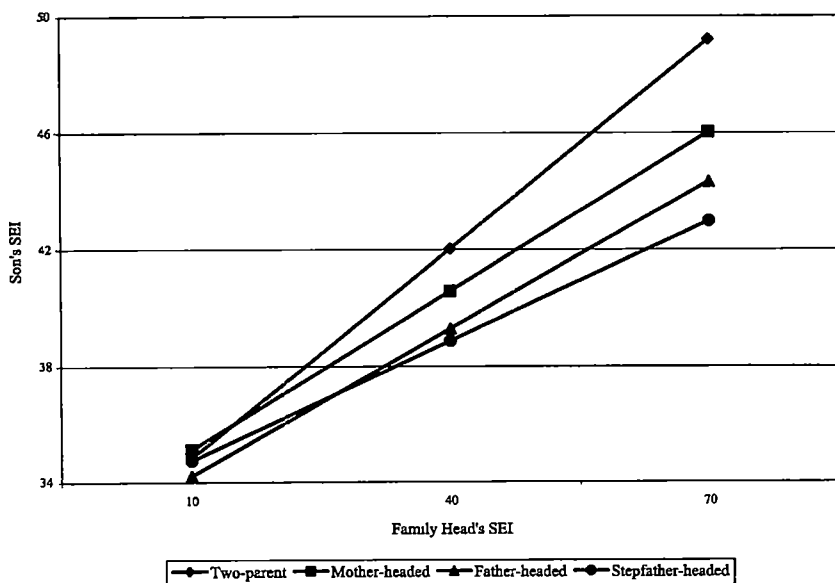


FIG. 1.—Estimated effects of childhood family structure and family head's SEI on son's current SEI.

origin SEI makes a difference of 14 SEI points for the son (from an SEI of 35 to 49), independent of all of the other variables in the model. In stepfamilies, moving from the low to the high end of origin SEI makes a difference of only 8 points in sons SEI (from 35 to 43). For single-mother families, the difference is 11 points, and for alternative father-headed families, 10 points.

This also means that family structure makes less of a difference for sons from low SES backgrounds than for sons from high SES backgrounds. At the low end of origin SEI, differences in destination SEI by family type are small. Children from low socioeconomic origins tend to end up in low socioeconomic destinations regardless of family type. Among children from the high end of origin SEI, differences in socioeconomic attainment across family types are substantial.[14]

We also speculated that children from some kinds of alternative families may have a more difficult time translating their educational achievements into occupational success because of more limited access to both parents' potential connections. To test this hypothesis, we followed steps like those above, adding son's education to the BIC-best model (table 5), and then terms representing interactions between each family type and son's education. All of the interactions were statistically significant at the 0.01 level. The results are displayed in figure 2.

Figure 2 shows a pattern that to our knowledge has not been documented empirically before: children from alternative families get less return (in terms of occupational SEI) to their educations, relative to children from two-parent homes. The differences look relatively small visually. This is somewhat deceptive because education is a hugely important determinant of occupational status for all family types (i.e., all the lines are steep).

Independent of other factors (including origin SEI, period, and interactions involving period), children from two-parent families gain, on average, 4 points on the SEI for each additional year of education that they obtain. Children from stepfamilies gain only 3 SEI points for each addi-

---

[14] The interaction effects displayed in figure 1 are adjusted for all the independent variables. As with the main effect, the interactive effect of single-mother families depends somewhat on the mix of control variables. In additional analysis (not shown, available from authors) we reestimated the interaction model but with only the controls shown in model 3 of table 4, plus period effects. The interaction effect continued to obtain (i.e., a less steep slope for children from single-mother families relative to those from two-biological-parent families, suggesting a weaker level of intergenerational transmission of socioeconomic status) but with smaller distances between points on the two slopes, suggesting a much smaller "direct" negative effect of single-mother families. Consistent with other findings throughout the article, the patterns for respondents from alternative male-headed families continued to be distinct from those of respondents from both single-mother and two-biological-parent families.
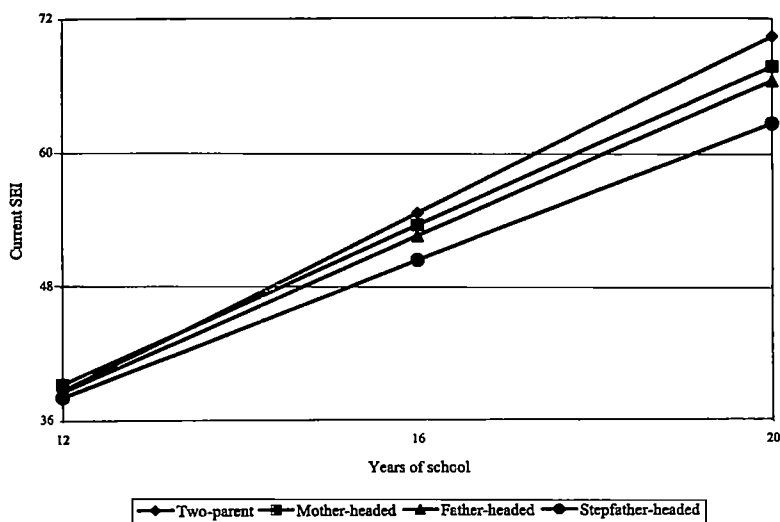
FIG. 2.—Estimated effects of childhood family structure and education on current SEI.

tional year of school. Children from single-mother and single-father families average approximately a half SEI point less (3.5) for each additional year of school than those from two-biological-parent families. At very low levels of education, children from alternative families are predicted to actually do better (in terms of SEI) than those from two-biological-parent families. But from the level of high school and beyond, family structure starts making an important difference to occupational returns in the opposite direction. This pattern is consistent with our idea that alternative families may be less able to connect their children into jobs commensurate with their educations. However, we predicted that the effect would be greatest for sons in single-mother families. It turns out that the effect is greatest for sons from stepfather-headed families.

## DISCUSSION

We went into the analysis guided by six questions. In the following discussion, we return to the questions and assess how the empirical analysis informs our understanding of each.

*Do conclusions about family structure effects change substantially depending on the mix of control variables?*

The answer to this question is yes. The analysis shows that the effect of growing up in a single-mother family is a complex function of a set of

factors that represent both risks and benefits to children's socioeconomic success. On the side of risk, relative to sons from two-biological-parent families, sons from single-mother families have (1) the disadvantage of having a family head with a greater average likelihood of unemployment and (2) the disadvantage of having a family head with a lower average occupational position. Across four large surveys spanning 30 years, across two dependent variables—children's education and occupation—we find that there is no effect of growing up in a single-mother family once family head's socioeconomic location (employment, occupation) is taken into account.

This finding may mean exactly what it appears on face—that parent's labor force attachment and occupational position are keys to understanding the effect of single-mother families on children's socioeconomic destinations. It directs us to explore job differences between single mothers and other family heads: Where are they located in the occupational structure? How do their job conditions affect the family? What are the employment opportunities and constraints (Glass and Camarigg 1992)? Why is parental employment good for children? In terms of policy, the findings suggest that adequate job opportunities for single mothers could go a long way toward diminishing the unfavorable consequences of single-mother families for children.

Another possibility is that these variables—parent's employment and occupation—are carrying the effects of other variables within them. One possibility is income. We may be tapping the level of economic resources available to families. Hauser and Warren (1997) entertain the possibility that the occupational index SEI may actually be a better measure of permanent income than income itself.

The findings dovetail in many ways with the work of Kiernan and colleagues (e.g., Kiernan 1997) in Great Britain. Based on the National Child Development Study, Kiernan (1997) found that differences in educational and work outcomes between children from two-biological-parent and single-mother homes are largely a function of financial hardships. She concluded that it may be the financial advantages held by children from two-biological-parent homes that propel them to higher attainments, and not necessarily the fact that their parents stayed together.

We also found that children from single-mother families benefited from a good average level of origin education and low average number of siblings. These features help to offset other kinds of disadvantages associated with single-mother families. If we hold constant origin education and siblings through statistical control, the negative effect of a single-mother family reemerges.

This pattern is most clearly evident in the analysis of educational attainment shown in the appendix (table A1). Across all the surveys, model 3 shows no effect of alternative mother-headed families, but model 5 does

show a significant negative effect of alternative mother-headed families. The difference in conclusion depends on the variables we decide to control statistically. Studies that treat parents' education and number of siblings as exogenous (e.g., Duncan 1967; McLanahan and Sandefur 1994) will actually begin with a kind of statistically heightened "total" effect of alternative mother-headed families. In turn, whether one embraces model 3- versus model 5-like results will shape family discourse in very different ways. The former model suggests that single mothers who are able to secure adequate positions in the social structure—indicated by their employment and occupation—can offset the negative effect of the loss of the father, and their children will do approximately as well (in education) as those from two-biological-parent families. The latter model suggests something closer to a "pathology of matriarchy" perspective—pointing to some long-term legacies of father absence for children's attainments that go beyond the loss of socioeconomic resources.

*Do conclusions about family structure effects change substantially depending upon period?*

The effects of alternative family structures on children's socioeconomic success have remained constant over the past 30 years. Change in the effect of family structure over time probably cannot explain discrepancies observed in the family structure effects literature.

The effects of other dimensions of family background have changed over the period. Race-based inequality in socioeconomic attainment is no longer on the decline—it has not been since 1987 or thereabouts—and it may even be increasing. Having a family head who was not employed became less detrimental to children over time, but this decline too stalled in the late 1980s and may be reversing course. While family head's SEI has declined fairly steadily as a determinant of children's success, family head's education has not. These patterns suggest that U.S. society may indeed be hitting a "speed bump" in its progress toward universalism (Hout 1997). The trends may have particularly grave implications for new cohorts from alternative families, who will be disproportionately more likely to face the triple threats of race, origin unemployment, and family structure.

*Do the socioeconomic achievements of children from different types of alternative families vary substantially?*

Because most studies compare the outcomes of children from two-biological-parent families with those of children from alternative families, we know less about differences or similarities in attainment among children from different kinds of alternative families. Our findings indicate that, over a 30-year period, children from single-mother families consistently do better than those raised in single-father families or stepfamilies, once socioeconomic position is taken into account (see also Amato and Keith 1991*b;* Amato and Booth 1991; Hoffmann and Johnson 1998).

Single-father families and stepfather-headed families have about the same negative effect on children's attainments, an effect that cannot be explained by socioeconomic position.[15]

The "pathology of matriarchy" position is that the lack of a man in the home would be particularly detrimental to boys. The sons from alternative male-headed households in our data lived either with their biological father or a stepfather, and they were able to report the male head's occupation. Averaging across the four surveys, male heads' occupations were about 4–5 points higher than those of the single mothers. Yet even with the socioeconomic advantage, these sons did no better than those from single-mother families, and in some cases they did worse. If the single mothers had occupied the same socioeconomic position as the single fathers and stepfathers, their children would have had significantly higher attainments than those from the alternative male-headed families.

That father presence gave no particular advantage to sons from alternative families may reflect the legacy of the death of the mother or of a neglectful mother. Stepfathers may have dual households to invest in, stepparent/child relations can be problematic, and stepfamilies may introduce some distance into the children's relationship with the mother (Biblarz et al. 1997). Whatever the explanation, none of them can be neatly construed to fit with a "pathology of matriarchy" view.

Prompted by Wojtkiewicz's (1993) conclusion that a dichotomous indicator of family type captures the main effects in most cases, Powell and Parcel (1997) adopted a two-parent family/not two-parent family measure, as did Biblarz and Raftery (1993). Other research treats single-mother families, single-father families, and stepfamilies together as a "one-parent" group (McLanahan and Sandefur 1994). Duncan and Duncan's (1969) "female-headed" families include both respondents from single-mother homes and those who reported living with a female head (e.g., grandmother, aunt) but not the mother. Single, female-headed families and single-parent families are not the same as single-mother families, even though the "single female" and "single parent" terms may evoke a "single mother" image. When all alternative female-headed or single-parent families are considered together, the large group of children from single-mother families gets combined with a smaller group of children from other types of alternative families in which the negative association with children's attainment is especially strong. This increases the negative effect associated with the combined group and so appears to reinforce a "pathology of matriarchy" perspective that may not be fully consistent with the data.

[15] The alternative father-headed families include single-father families but also those where the father had remarried. Unfortunately, the data did not allow us to distinguish these.

*Which (if any) of the theoretical perspectives are supported in terms of both the static and change predictions (table 1) that they make?*

Evolutionary parental investment theory was the only one where static and change predictions were both borne out by the data. The sociological theories faltered in that they provide no basis for our finding that children from single-mother families actually have some advantage over children from other kinds of alternative families. The static predictions of economic theory—that children from stepfamilies will do better than those from single-parent families, that children from single-father families will do better than those from single-mother homes—were entirely unsupported.

The stability over time in the effect of family structure provides some evidence for its not being a selection effect. Hence, the change predictions of the parental competence and marital conflict models are not supported. Approximately two-thirds of the OCG I sample from alternative families experienced the death of a parent. In contrast, two-thirds of the NSFH2 sample from alternative families experienced parental divorce. Selection models would suggest that the former group had a more competent (widowed) parent(s) or had less exposure to marital conflict than the latter group, and so should have done better. Yet the consequences of family structure were essentially the same for both groups.[16] The constancy of this effect may imply a fundamental family process resistant to changes in the times or the culture.

The evolutionary perspective differs from the others by placing gender of parent and biological relations at the forefront of an explanation of family structure effects. It predicts that children from single-mother homes will have advantages over those from single-father homes because mothers have more of their reproductive investment tied up in their children than fathers. This prediction was supported in that, holding constant other variables, children from single-mother homes had higher attainments than those from alternative father-headed households. The evolutionary perspective also predicts that a stepparent will be of no advantage to children (stepparents have no real incentive to invest in stepchildren since stepchildren contribute

---

[16] While the marital conflict model is not supported by our data, further testing is required. Assume that in the earlier period only the most conflicted marriages ended in divorce, whereas in the later period families that are unhappy but less so also divorce. It is possible, e.g., that in the earlier period the pool of respondents from alternative families was made up of a large proportion that experienced parental death and a small proportion that experienced "severe" divorce of parents. In the later period, the pool from alternative families may have been composed of a small proportion that experienced parental death and a large proportion that experienced less severe parental divorce. Divorce may become less severe also as it becomes more institutionalized (Cherlin 1978; Wolfinger 1998). These may be offsetting effects resulting in no change in the effect of family structure over time.

nothing to stepparents' fitness) and may actually represent a negative effect insofar as the stepparent competes with the children for the resources of the biological parent. This prediction was supported insofar as children from single-mother families had higher attainments than those from stepfamilies. Finally, evolutionary theory predicts no change over time in the magnitude of the effect of family structure. Parental investment determines children's outcomes. The experience of the death of a spouse versus divorce would not alter the custodial parents' incentive to invest in the children. Parental investment is a function of a reproductive calculus that is part of an evolved psychology developed during the foraging eons. This prediction of no change over time was also supported.

*Do alternative families exhibit a weaker level of intergenerational socioeconomic transmission than two-biological-parent families?*

*Do alternative families reduce children's ability to translate educational achievement into occupational success?*

The finding observed in OCG II—that the effect of origin SEI on destination SEI is weaker among children from alternative families (Biblarz and Raftery 1993)—was confirmed here using data that were both older and more recent, and that included both men and women. This supports the proposition that alternative family forms may produce a general weakening in the intergenerational transmission of socioeconomic position. The particularly interesting aspect of the finding in the present case is that family structure may be more important for children coming out of the higher than the lower end of the socioeconomic spectrum. Loosely, among children from blue-collar families, there are almost no differences in average socioeconomic attainment depending on whether one was raised in a two-parent family, a single-parent family, and so on. But there are noticeable differences in attainment by family type among children from middle and upper-middle class origins. There has been some suggestion in media and policy debates that family structure probably will not matter much for children coming out of high-resource, female-headed families. Our results suggest the opposite: it may be at the high end where family structure does matter.

In a similar way, we found that children from alternative families get less occupational return to higher education. While we have no direct measure, our speculation is that the sort of "favoritism in hiring and special favors" given to families to help them launch their children, of the sort detected by Hout (1989, p. 322) in the case of Ireland, may be more accessible to two-parent families than to alternative families. Interestingly, both of these interaction effects, one involving intergenerational transmission and the other involving intragenerational attainment over the life course, expose consequences of family structure for children that would be hidden in a more typical "direct" effects approach.

APPENDIX

## TABLE A1

UNSTANDARDIZED BETAS FROM REGRESSION OF YEARS OF SCHOOL ON SELECTED DIMENSIONS OF FAMILY BACKGROUND: U.S.-BORN RESPONDENTS AGES 25–64, FROM NONFARM ORIGINS

| Independent Variable | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1962 OCG I (N = 11,697): | | | | | |
| Childhood family structure: | | | | | |
| Two-biological-parent (reference) | ... | ... | ... | ... | ... |
| Alternative mother-headed | -.89*** | -.15 | -.04 | -.25 | -.45*** |
| Alternative father-headed | -.96*** | -.92*** | -.75*** | -.73*** | -.69*** |
| Alternative stepfather-headed | -.75** | -.56 | -.38 | -.37 | -.51 |
| Race (African-American = 1) | -2.30*** | -2.04*** | -1.20*** | -1.12*** | -1.02*** |
| Socioeconomic origins: | | | | | |
| Family head was not employed | | -1.72*** | -1.73*** | -1.26*** | -1.11*** |
| Family head's occupational SEI | | | .06*** | .04*** | .04*** |
| Family head's education | | | | .59*** | .51*** |
| Family head's education not reported | | | | -1.40*** | -1.27*** |
| Number of siblings | | | | | -.26*** |
| Intercept | 11.66*** | 11.73*** | 9.74*** | 9.40*** | 10.69*** |
| $R^2$ | .0457 | .0644 | .1952 | .2458 | .2851 |
| BIC | -510 | -732 | -2,484 | -3,225 | -3,841 |

1973 OCG II (N = 19,283):

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Childhood family structure:** | | | | | |
| Two-biological-parent (reference) | ... | ... | ... | ... | ... |
| Alternative mother-headed | −.80*** | −.54*** | −.25* | −.31*** | −.44*** |
| Alternative father-headed | −1.27*** | −1.22*** | −.96*** | −.87*** | −.80*** |
| Alternative stepfather-headed | −1.06*** | −.99*** | −.90*** | −.86*** | −.87*** |
| Race (African-American = 1) | −1.37*** | −1.36*** | −.59*** | −.60*** | −.39*** |
| **Socioeconomic origins:** | | | | | |
| Family head was not employed | | −1.12*** | −1.12*** | −.80*** | −.61*** |
| Family head's occupational SEI | | | .05*** | .03*** | .03*** |
| Family head's education | | | | .53** | .45*** |
| Family head's education not reported | | | | −1.45*** | −1.33*** |
| Number of siblings | | | | | −.25*** |
| Intercept | 12.55 | 12.60 | 10.88 | 10.50 | 11.64 |
| $R^2$ | .0319 | .0409 | .1697 | .2223 | .2649 |
| BIC | −586 | −756 | −3,527 | −4,769 | −5,846 |

1986–88 SIPPs (N = 34,825):

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Childhood family structure:** | | | | | |
| Two-biological-parent (reference) | ... | ... | ... | ... | ... |
| Alternative mother-headed | −.47*** | −.26*** | .12 | −.26*** | −.24*** |
| Alternative father-headed | −.59*** | −.56*** | −.52*** | −.42*** | −.36*** |
| Alternative stepfather-headed | −.68*** | −.64*** | −.54*** | −.47*** | −.45*** |
| Race (African-American = 1) | −1.09*** | −1.06*** | −.62*** | −.38*** | −.15 |
| Gender (female = 1) | −.21*** | −.21*** | −.22*** | −.24*** | −.23*** |
| African-American × female | .54*** | .54*** | .56*** | .50*** | .49*** |
| **Socioeconomic origins:** | | | | | |
| Family head was not employed | | −.73*** | −.77*** | −.07 | −.26*** |
| Family head's occupational SEI | | | .07*** | .03*** | .03*** |

359

TABLE A1 (Continued)

| Independent Variable | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Family head's education | | | | .50*** | .45*** |
| Family head's education not reported | | | | −1.20*** | −1.17*** |
| Number of siblings | | | | | −.18*** |
| Intercept | 13.23 | 13.28 | 10.90 | 10.53 | 11.34 |
| $R^2$ | .0227 | .0304 | .1342 | .2007 | .2308 |
| BIC | −737 | −1,002 | −4,935 | −7,697 | −9,023 |
| **1992–94 NSFH2 ($N = 6,536$):** | | | | | |
| Childhood family structure: | | | | | |
| Two-biological-parent (reference) | ... | ... | ... | ... | ... |
| Alternative mother-headed | −.49*** | −.42*** | −.11 | −.39*** | −.42*** |
| Alternative father-headed | −.70*** | −.68*** | −.68*** | −.60*** | −.41* |
| Alternative stepfather-headed | −.79*** | −.74*** | −.65*** | −.57*** | −.57*** |
| Race (African-American = 1) | −1.40*** | −1.36*** | −.96*** | −.62*** | −.45** |
| Gender (female = 1) | −.35*** | −.35*** | −.36*** | −.30*** | −.29*** |
| African-American × female | .84*** | .85*** | .88*** | .84*** | .88*** |
| Socioeconomic origins: | | | | | |
| Family head was not employed | | −.62*** | −.66*** | −.26*** | −.20 |
| Family head's occupational SEI | | | .06*** | .03*** | .03*** |
| Family head's education | | | | .38*** | .34*** |
| Family head's education not reported | | | | −1.18*** | −1.10*** |
| Number of siblings | | | | | −.13*** |
| Intercept | 13.66 | 13.70 | 11.52 | 10 94 | 11.53 |
| $R^2$ | .0418 | .0472 | .1540 | .2402 | .2602 |
| BIC | −226 | −255 | −1,023 | −1,708 | −1,873 |

\* $|t| \geq \sqrt{\ln N + 2}$
\** $|t| \geq \sqrt{\ln N + 6}$
\*** $|t| \geq \sqrt{\ln N + 10}$

# REFERENCES

Acock, Alan C., and David H. Demo. 1994. *Family Diversity and Well-Being.* Thousand Oaks, Calif.: Sage.

Acock, Alan C., and K. Jill Keicolt. 1989. "Is It Family Structure or Socioeconomic Status? Family Structure during Adolescence and Adult Adjustment." *Social Forces* 68:553–71.

Amato, Paul R., and Alan Booth. 1991. "Consequences of Parental Divorce and Marital Unhappiness for Adult Well-Being." *Social Forces* 69:895–914.

———. 1997. *A Generation at Risk: Growing Up in an Era of Family Upheaval.* Cambridge, Mass.: Harvard University Press

Amato, Paul R., and Bruce Keith. 1991a. "Parental Divorce and Adult Well-Being: A Meta-Analysis." *Journal of Marriage and the Family* 53:43–58.

———. 1991b. "Separation from a Parent during Childhood and Adult Socioeconomic Attainment." *Social Forces* 70:187–206.

Amato, Paul R., Laura Spencer Loomis, and Alan Booth. 1995. "Parental Divorce, Marital Conflict, and Offspring Well-Being during Early Adulthood." *Social Forces* 73:895–915.

Astone, Nan Marie, and Sara S. McLanahan. 1991. "Family Structure, Parental Practices, and High School Completion." *American Sociological Review* 56:309–20.

Baumrind, Diana. 1978. "Parental Disciplinary Patterns and Social Competence in Children." *Youth and Society* 9:239–75.

———. 1980. "New Directions in Socialization Research." *American Psychologist* 35: 639–52.

Becker, Gary S. 1964. *Human Capital.* New York: National Bureau of Economic Research.

———. 1981. *A Treatise on the Family.* Cambridge, Mass.: Harvard University Press.

Becker, Gary S., and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics* 4:S1–S39.

Biblarz, Timothy J., and Greg Gottainer. 1999. "Single-Mother Family Structures Produced by Death of Spouse and Divorce: Why Do They Have Different Consequences for Children?" *Journal of Marriage and the Family,* in press.

Biblarz, Timothy J., and Adrian E. Raftery. 1993. "The Effects of Family Disruption on Social Mobility." *American Sociological Review* 58:97–109.

Biblarz, Timothy J., Adrian E. Raftery, and Alexander Bucur. 1997. "Family Structure and Social Mobility." *Social Forces* 75:1319–39.

Blake, Judith. 1989. *Family Size and Achievement.* Berkeley, Calif.: University of California Press.

Blau, Peter M., and Otis Dudley Duncan. 1967. *The American Occupational Structure.* New York: Wiley.

Boggess, Scott. 1998. "Family Structure, Economic Status, and Educational Attainment." *Journal of Population Economics* 11:203–22.

Bronfenbrenner, Urie. 1979. *The Ecology of Human Development: Experiments by Nature and Design.* Cambridge, Mass.: Harvard University Press.

Bumpass, Larry L., and James A. Sweet. 1989. "Children's Experience in Single-Parent Families: Implications of Cohabitation and Marital Transitions." *Family Planning Perspectives* 21:256–60.

Cherlin, Andrew. 1978. "Remarriage as an Incomplete Institution." *American Journal of Sociology* 84:634–50.

Cherlin, Andrew J., P. Lindsay Chase-Lansdale, and Christine McRae. 1998. "Effects of Parental Divorce on Mental Health throughout the Life Course." *American Sociological Review* 63:239–49

Cherlin, Andrew J., and Frank F. Furstenberg, Jr. 1994. "Stepfamilies in the United States: A Reconsideration." *Annual Review of Sociology* 20:359–81.

Cherlin, Andrew J., Frank F. Furstenberg, P. Lindsay Chase-Lansdale, Kathleen Kiernan, Donna Ruane Morrison, and Julien Teitler. 1991. "Longitudinal Studies of Effects of Divorce on Children in Great Britain and the United States." *Science* 252:1386–89.

Crosby, Faye J., ed. 1987. *Spouse, Parent, Worker: On Gender and Multiple Roles.* New Haven, Conn.· Yale University Press.

Daly, Martin, and Margo I. Wilson. 1996 "Violence against Stepchildren." *Current Directions in Psychological Science* 5:77–81.

DaVanzo, Julie, and M. Omar Rahman. 1993. "American Families: Trends and Correlates." *Population Index* 59:350–86.

Dawson, Deborah A. 1991. "Family Structure and Children's Health and Well-Being: Data from the 1988 National Health Interview Survey on Child Health." *Journal of Marriage and the Family* 53:573–84.

DiPrete, Thomas A., and David B. Grusky. 1990. "Structure and Trend in the Process of Stratification for American Men and Women." *American Journal of Sociology* 96:107–43.

Duncan, Beverly. 1967. "Education and Social Background." *American Journal of Sociology* 72:363–72.

Duncan, Beverly, and Otis Dudley Duncan. 1969. "Family Stability and Occupational Success." *Social Problems* 16:273–85.

Duncan, Otis Dudley. 1961. "A Socioeconomic Index for All Occupations." Pp. 109–38 in *Occupations and Social Status,* edited by A. J. Reiss Jr. Glencoe, Ill.: Free Press.

Emlen, Stephen T 1997. "The Evolutionary Study of Human Family Systems." *Social Science Information Sur Les Sciences Sociales* 36:563–89.

Featherman, David L., and Robert M. Hauser. 1978. *Opportunity and Change.* New York: Academic Press.

Furstenberg, Frank F, Jr., and Andrew J. Cherlin. 1991. *Divided Families.* Cambridge, Mass.: Harvard University Press.

Glass, Jennifer, and Valerie Camarigg. 1992. "Gender, Parenthood, and Job-Family Compatibility." *American Journal of Sociology* 98:131–51.

Glenn, Norval D. 1994 "The Re-evaluation of Family Change by American Social Scientists." Paper presented to the Committee for the International Year of the Family of the Catholic Archdiocese of Melbourne, Melbourne.

Glenn, Norval D., and Kathryn B. Kramer. 1987. "The Marriages and Divorces of the Children of Divorce." *Journal of Marriage and the Family* 49:811–25.

Grusky, David B., and Thomas A. DiPrete. 1990. "Recent Trends in the Process of Stratification " *Demography* 27:617–38.

Grych, John H., and Frank D. Fincham. 1990. "Marital Conflict and Children's Adjustment: A Cognitive-Contextual Framework." *Psychological Bulletin* 108:267–90.

Hauser, Robert M., and John Robert Warren. 1997. "Socioeconomic Indexes for Occupations: A Review, Update, and Critique." *Sociological Methodology* 27:177–298.

Hetherington, E. Mavis, Martha Cox, and Roger Cox. 1978. "The Aftermath of Divorce." Pp. 149–76 in *Mother-Child Father-Child Relationships,* edited by J. H. Stevens Jr. and M. Mathews. Washington, D.C.: National Association for the Education of Young People.

Hoffmann, John P., and Robert A Johnson. 1998. "A National Portrait of Family Structure and Adolescent Drug Use." *Journal of Marriage and the Family* 60:633–45.

Holden, Karen C., and Pamela J. Smock. 1991. "The Economic Costs of Marital Disso-

lution: Why Do Women Bear a Disproportionate Cost?" *Annual Review of Sociology* 17:51–78.

Hout, Michael. 1984a. "Occupational Mobility of Black Men in the United States: 1962 to 1973." *American Sociological Review* 49:308–22.

———. 1984b. "Status, Autonomy, and Training in Occupational Mobility." *American Journal of Sociology* 89:1397–1409

———. 1988. "More Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s." *American Journal of Sociology* 93:1358–400.

———. 1989. *Following in Father's Footsteps: Social Mobility in Ireland.* Cambridge, Mass.: Harvard University Press.

———. 1997. "Speed Bumps on the Road to Meritocracy: Occupational Mobility of Women and Men in the United States, 1972–1994." Paper presented in the Lectures on Causality in the Social Sciences Series, University of Washington, Department of Sociology, Spring.

Kiernan, Kathleen. 1997. "The Legacy of Parental Divorce: Social, Economic and Demographic Experiences in Adulthood " Working paper. London School of Economics, Centre for Analysis of Social Exclusion.

Kline, Marsha, Janet R. Johnston, and Jeanne M. Tschann. 1991 "The Long Shadow of Marital Conflict: A Model of Children's Postdivorce Adjustment." *Journal of Marriage and the Family* 53:297–310.

Kohn, Melvin L. 1969. *Class and Conformity: A Study in Values,* 2d ed. Chicago: University of Chicago Press.

———. 1983. "On the Transmission of Values in the Family: A Preliminary Formulation." *Research in Sociology of Education and Socialization* 4:1–12.

Logan, John Allen. 1996. "Opportunity and Choice in Socially Structured Labor Markets." *American Journal of Sociology* 102:114–60.

London, Rebecca A. 1996 "The Difference between Divorced and Never-Married Mothers' Participation in the Aid to Families with Dependent Children Program." *Journal of Family Issues* 12:170–85

Mauldon, Jane. 1990. "The Effect of Marital Disruption on Children's Health." *Demography* 27:431–46

McLanahan, Sara S. 1985. "Family Structure and the Reproduction of Poverty." *American Journal of Sociology* 90·873–901.

McLanahan, Sara S., Irwin Garfinkel, and Theodora Ooms. 1987. "Female-Headed Families and Economic Policy: Expanding the Clinician's Focus." Pp. 1–18 in *Clinical Issues in Single-Parent Households,* edited by J. C. Hansen and M Lindblad-Goldberg. Rockville, Md.: Aspen Publishers.

McLanahan, Sara, and Gary Sandefur. 1994. *Growing Up with a Single Parent.* Cambridge, Mass.: Harvard University Press.

Meyer, Daniel R., and Steven Garasky 1993. "Custodial Fathers: Myths, Realities, and Child Support Policies." *Journal of Marriage and the Family* 55:73–90.

Moss, Philip, and Chris Tilly. 1996. "Raised Hurdles for Black Men: Evidence from Interviews with Employers." Working paper. Russell Sage Foundation.

Mott, Frank L., Lori Kowaleski-Jones, and Elizabeth G. Menaghan. 1997. "Paternal Absence and Child Behavior: Does a Child's Gender Make a Difference?" *Journal of Marriage and the Family* 59:103–18.

Moynihan, Daniel Patrick. 1965. *The Negro Family: The Case for National Action.* Washington, D.C.: U.S. Department of Labor.

Mueller, Charles W., and Hallowell Pope. 1977. "Marital Instability: A Study of Its Transmission between Generations." *Journal of Marriage and the Family* 39:83–93.

Nock, Steven L. 1988. "The Family and Hierarchy." *Journal of Marriage and the Family* 50:957–66.

Painter, Gary. 1998. "Family Structure and the Achievement of Youths: The Impor-

tance of Parental Involvement in Education." Working paper. University of Southern California, School of Public Administration.

Parcel, Toby L., and Elizabeth G. Menaghan. 1994. *Parents' Jobs and Children's Lives.* Hawthorne, N.Y.: Aldine DeGruyter.

Popenoe, David. 1988. *Disturbing the Nest: Family Change and Decline in Modern Societies.* New York: Aldine DeGruyter.

———. 1993. "American Family Decline, 1960–1990: A Review and Appraisal." *Journal of Marriage and the Family* 55:527–41.

———. 1996. *Life without Father.* New York: Free Press.

Powell, Brian, and Lala Carr Steelman. 1993. "The Educational Benefits of Being Spaced Out. Sibship Density and Educational Progress." *American Sociological Review* 58:367–81.

Powell, Mary Ann, and Toby L. Parcel. 1997. "Effects of Family Structure on the Earnings Attainment Process: Differences by Gender." *Journal of Marriage and the Family* 59:419–33.

Raftery, Adrian E. 1986a. "A Note on Bayes Factors for Log-Linear Contingency Table Models with Vague Prior Information." *Journal of the Royal Statistical Society,* series B, 48:249–50.

———. 1986b. "Choosing Models for Cross-Classifications." *American Sociological Review* 51:145–46.

———. 1995. "Bayesian Model Selection in Social Research (with Discussion)." Pp. 111–95 in *Sociological Methodology 1995,* edited by Peter V. Marsden. Cambridge, Mass.: Blackwell Publishers.

Rozondal, Frederick G. "Halos vs. Stigmas: Long-Term Effects of Parent's Death or Divorce on College Students' Concepts of the Family." *Adolescence* 18:947–55.

Silverman, Phyllis R., Steven Nickman, and J. William Worden. 1992. "Detachment Revisited: The Child's Reconstruction of a Dead Parent." *American Journal of Orthopsychiatry* 62:494–503.

Simons, Ronald L., and Associates. 1996. *Understanding Differences between Divorced and Intact Families: Stress, Interactions, and Child Outcome.* Thousand Oaks, Calif.: Sage.

Stacey, Judith. 1996. *In the Name of the Family: Rethinking Family Values in the Postmodern Age.* Boston: Beacon Press.

Steelman, Lala Carr, and Brian Powell. 1989. "Acquiring Capital for College: The Constraints of Family Configuration." *American Sociological Review* 54:844–55.

Sugarman, Stephen. 1993. "Reforming Welfare through Social Security." *University of Michigan Journal of Law Reform* 26:849–51.

———. 1995. "Financial Support of Children and the End of Welfare as We Know It." *Virginia Law Review* 81:2523–73.

Teachman, Jay D., Kathleen Paasch, and Karen Carver. 1997. "Social Capital and the Generation of Human Capital." *Social Forces* 75:1343–60.

Thomson, Elizabeth, Thomas L. Hanson, and Sara S. McLanahan. 1994. "Family Structure and Child Well-Being: Economic Resources vs. Parental Behaviors." *Social Forces* 73:221–42.

Thomson, Elizabeth, Sara S. McLanahan, and Roberta Braus Curtin. 1992. "Family Structure, Gender, and Parental Socialization." *Journal of Marriage and the Family* 54:368–78.

Trivers, Robert. 1972. "Parental Investment and Sexual Selection." Pp. 136–79 in *Sexual Selection and the Descent of Man,* edited by Bernard Campbell. Chicago: Aldine DeGruyter.

Wallerstein, Judith S. 1989. *Second Chances: Men, Women and Children a Decade after Divorce.* New York: Ticknor and Fields.

Weinstein, Maxine, and Arland Thornton. 1989. "Mother-Child Relations and Adolescent Sexual Attitudes and Behavior." *Demography* 26:563–77.

Wojtkiewicz, Roger A. 1993. "Simplicity and Complexity in the Effects of Parental Structure on High School Graduation." *Demography* 30:701–17.

Wojtkiewicz, Roger A., Sara S. McLanahan, and Irwin Garfinkel. 1990. "The Growth of Families Headed by Women, 1950–1980." *Demography* 27:19–30.

Wolfinger, Nicholas H. 1998. "Trends in the Intergenerational Transmission of Divorce." Paper presented at the 1996 meetings of the Society for the Study of Social Problems, and the American Sociological Association, New York, August.

Wu, Lawrence L., and Brian C. Martinson. 1993. "Family Structure and the Risk of a Premarital Birth." *American Sociological Review* 58:210–32.

# Marx's Theory of Metabolic Rift: Classical Foundations for Environmental Sociology[1]

John Bellamy Foster
*University of Oregon*

This article addresses a paradox: on the one hand, environmental sociology, as currently developed, is closely associated with the thesis that the classical sociological tradition is devoid of systematic insights into environmental problems; on the other hand, evidence of crucial classical contributions in this area, particularly in Marx, but also in Weber, Durkheim, and others, is too abundant to be convincingly denied. The nature of this paradox, its origins, and the means of transcending it are illustrated primarily through an analysis of Marx's theory of metabolic rift, which, it is contended, offers important classical foundations for environmental sociology.

## CLASSICAL BARRIERS TO ENVIRONMENTAL SOCIOLOGY

In recent decades, we have witnessed a significant transformation in social thought as various disciplines have sought to incorporate ecological awareness into their core paradigms in response to the challenge raised by environmentalism and by what is now widely perceived as a global ecological crisis. This transformation has involved a twofold process of rejecting much of previous thought as ecologically unsound, together with an attempt to build on the past, where possible. This can be seen as occurring with unequal degrees of success in the various disciplines. Geography, with its long history of focusing on the development of the natural landscape and on biogeography (see Sauer 1963), was the social science that adapted most easily to growing environmental concerns. Anthropol-

ogy, with a tradition of investigating cultural survival and its relation to ecological conditions (see Geertz 1963; Milton 1996), also adjusted quickly to a period of greater environmental awareness. In other social science disciplines, significant progress in incorporating ecological ideas has been made, yet with less discernible effect on the core understandings of these fields. Economics, which was able to draw on the theoretical foundations provided by A. C. Pigou's *Economics of Welfare* (1920), has seen the rapid development of a distinctive, if limited, approach to environmental issues focusing on the internalization of "externalities"—making "environmental economics . . . one of the fastest-growing academic sub-disciplines throughout the industrial world" (Jacobs 1994, p. 67). As a relatively atheoretical field, political science has had little difficulty in incorporating environmental issues into its analysis of public policy, its focus on pluralist interest groups, its social contract theory, and more recently its emphasis on rational choice (Dryzek 1997)—though the pragmatic character of most political science in the United States, together with the lack of a strong Green political party and the absence of a clear connection between identification with environmental causes and voting behavior, has kept the politics of the environment on the margins of the discipline.

In sociology too, dramatic progress has been made, as seen by the rapid growth of the subfield of environmental sociology in the 1970s and again (after a period of quiescence) in the late 1980s and 1990s (see Dunlap 1997). Nevertheless, sociology is perhaps unique within the social sciences in the degree of resistance to environmental issues. An early barrier erected between society and nature, sociology and biology—dividing the classical sociologies of Marx, Weber, and Durkheim from the biological and naturalistic concerns that played a central role in the preclassical sociology of the social Darwinists—has hindered the incorporation of environmental sociology within the mainstream of the discipline, according to an interpretation repeatedly voiced by prominent environmental sociologists over the last two and a half decades (Burch 1971, pp. 14–20; Dunlap and Catton 1979, pp. 58–59; Benton 1994, pp. 28–30; Murphy 1994, pp. ix–x; Beck 1995, pp. 117–20; Buttel 1996, pp. 57–58; Murphy 1996).

Hence, until recently "there has . . . been general agreement among environmental sociologists that the classical sociological tradition has been inhospitable to the nurturing of ecologically-informed sociological theory" (Buttel 1986, p. 338). "From an environmental-sociological point of view," Buttel (1996, p. 57) has argued, "the classical tradition can be said to be 'radically sociological,' in that in their quest to liberate social thought and sociology from reductionisms, prejudices, power relations, and magic, the classical theorists (and, arguably more so, the 20th century interpreters of the classical tradition) wound up exaggerating the auton-

omy of social processes from the natural world." Likewise, Benton (1994, p. 29) has observed that "the conceptual structure or 'disciplinary matrix' by which sociology came to define itself, especially in relation to potentially competing disciplines such as biology and psychology, effectively excluded or forced to the margins of the discipline questions about the relations between society and its 'natural' or 'material' substrate." "Sociology," according to one prominent environmental sociologist, "was constructed as if nature didn't matter" (Murphy 1996, p. 10). Such marginalization of the physical environment was made possible, in part, through the enormous economic and technological successes of the industrial revolution, which have long given the impression that human society is independent of its natural environment (Dunlap and Martin 1983, pp. 202–3). This is seen as offering an explanation for the fact that "sociological work on resource scarcity never appeared in the discipline's top journals" in the United States (Dunlap 1997, p. 23; also Dunlap and Catton 1994, p. 8).

Modern sociology in its classical period, according to the prevailing outlook within environmental sociology, was consolidated around a humanistic worldview that emphasized human distinctiveness in relation to nature. This has been referred to by some as the old "human exemptionalist paradigm" in contrast to the "new environmental paradigm," which rejects the anthropocentrism supposedly characteristic of the former view (Catton and Dunlap 1978; Dunlap and Catton 1994). With respect to Durkheim, for example, it has been argued that the social constituted a distinct reality, relatively autonomous from the physical individual and from psychological and biological pressures (Benton and Redclift 1994, p. 3; Dunlap and Catton 1979, p. 58). "The thrust of Durkheim's and Weber's methodological arguments," according to Goldblatt (1996, p. 3), was to cordon off sociology from biology and nature, rejecting "all forms of biological determinism"; while Marx's treatment of such issues, though considerable, was largely confined to the "marginal" realm of agricultural economics.

In the language of contemporary environmentalism, then, sociology is a discipline that is "anthropocentric" in orientation, allowing little room for consideration of society's relation to nature, much less the thoroughgoing "ecocentrism" proposed by many environmentalists. It is rooted in a "socio-cultural determinism" that effectively excludes ecological issues (Dunlap and Martin 1983, p. 204). For Dunlap and Catton (1994, p. 6), sociology needs to shed "the 'blinders' imposed by [human] exceptionalism" and to acknowledge "the ecosystem dependence of all human societies."

One result of this problem of theoretical dissonance is that environmental sociology, despite important innovations, has continued to have only a

marginal role within the discipline as a whole. Although an environmental sociology section of the American Sociological Association was launched in 1976, it did not have the paradigm-shifting effect on sociology that leading figures in the section expected. Neither was sociology as a whole much affected by the rise of environmental sociology, nor did environmental issues gain much notice within the profession. As one leading practitioner of environmental sociology observed in 1987, "The discipline at large has handily withstood the challenges to its theoretical assumptions posed by environmental sociologists" (Buttel 1987, p. 466).

Where the core sociological discipline has been most ready to acknowledge environmental issues is in the area of environmental movements. There the literature has rapidly expanded in recent years through the growth of the environmental justice movement, concerned with the impact of environmental degradation on distinct sociological groupings, conceived in terms of race, class, gender, and international hierarchy. But this literature owes much more to social movement theory than to the environmentalist challenge to traditional sociological conceptions.

One way in which environmental sociologists have sought to address this problem of what are generally perceived as barriers within classical sociology to any consideration of the physical environment is by reaching out to the preclassical social Darwinist tradition: thinkers such as Malthus and Sumner (Catton 1982). Recently, however, there has been a great deal of research within environmental sociology directed not at circumventing the main classical sociological theorists but at unearthing alternative foundations within the classical literature, neglected in later interpretations. For example, an impressive attempt has been made by Murphy (1994) to establish a neo-Weberian sociology by applying Weber's critique of rationalization to the ecological realm and developing an "ecology of social action." Järvikoski (1996) has argued that we should reject the view that Durkheim simply neglected nature, choosing to address instead Durkheim's social constructionism with respect to nature, while examining how society fit within the hierarchical conception of nature that he generally envisioned. Others have stressed Durkheim's use of biological analogies and the demographic basis that he gave to his social morphology of the division of labor and urbanism, which seemed to foreshadow the urban-oriented human ecology of Park and other Chicago sociologists (Buttel 1986, pp. 341–42). The most dramatic growth of literature in relation to classical sociology, however, has centered on Marx's ecological contributions, which were more extensive than in the other classical theorists, and which have spawned a vast and many-sided international debate, encompassing all stages of Marx's work (e.g., Schmidt 1971; Parsons 1977; Giddens 1981; Redclift 1984; Clark 1989; Benton 1989; McLaughlin 1990;

Mayumi 1991; Grundmann 1991; Eckersley 1992; Perelman 1993; Hayward 1994; Harvey 1996; Burkett 1997; Foster 1997; Dickens 1997; O'Connor 1998).

Significantly, this growing literature on the relation of classical sociological theorists to environmental analysis has caused some of the original critics of classical sociology within environmental sociology to soften their criticisms. Buttel, one of the founders of the subdiscipline, has gone so far as to suggest that, despite all of their deficiencies in this respect, "a meaningful environmental sociology can be fashioned from the works of the three classical theorists" (1986, pp. 340–41). We now know, for example, that Weber, writing as early as 1909 in his critique of Wilhelm Ostwald's social energetics, demonstrated some concern over the continued availability of scarce natural resources and anticipated the ecological economist Georgescu-Roegen in arguing that the entropy law applied to materials as well as energy (Martinez-Alier 1987, pp. 183–92). Durkheim's analysis of the implications of Darwinian evolutionary theory—as we shall see below—pointed toward a complex, coevolutionary perspective. Nevertheless, the widespread impression of rigid classical barriers to environmental sociology continues to exert its influence on most environmental sociologists, leaving them somewhat in the state of the mythical centaur, with the head of one creature and the body of another, unable fully to reconcile their theoretical commitment to classical sociology with their environmental sociology, which demands that an emphasis be placed on the relations between society and the natural environment.

The following will focus on addressing the seemingly paradoxical relation of classical sociological theory and environmental sociology by centering on the work of Marx, while referring only tangentially to the cases of Weber and Durkheim. It will be argued that neglected but crucial elements within Marx's social theory offer firm foundations for the development of a strong environmental sociology. In contrast to most treatments of Marx's ecological writings, emphasis will be placed not on his early philosophical works but rather on his later political economy. It is in the latter that Marx provided his systematic treatment of such issues as soil fertility, organic recycling, and sustainability in response to the investigations of the great German chemist Justus von Liebig—and in which we find the larger conceptual framework, emphasizing the metabolic rift between human production and its natural conditions.[2]

It may seem ironic, given Marx's peculiar dual status as an insider-founder and outsider-critic of classical sociology (not to mention his repu-

[2] The issue of sustainability, or the notion that basic ecological conditions need to be maintained so that the ability of future generations to fulfill their needs will not be compromised, is the leitmotif of most contemporary environmental thought.

tation in some quarters as an enemy of nature), to turn to him in order to help rescue sociology from the embarrassing dilemma of having paid insufficient attention to the relation between nature and human society. Yet, the discovery or rediscovery of previously neglected features of Marx's vast intellectual corpus has served in the past to revitalize sociology in relation to such critical issues as alienation, the labor process, and, more recently, globalization. The irony may seem less, in fact, when one considers that there already exists "a vast neo-Marxist literature in environmental sociology, and [that] there are few other areas of sociology today that remain so strongly influenced by Marxism" (Buttel 1996, p. 61).

In constructing this argument around Marx, an attempt will be made to comment more broadly on the paradox of the existence—as we are now discovering—of a rich body of material on environmental issues within classical sociological theory, on the one hand, and the widespread perception that the classical tradition excluded any serious consideration of these issues, and itself constitutes a barrier inhibiting the development of environmental sociology, on the other. Here two hypotheses will be advanced arising out of the treatment of Marx. First, the apparent blindness of classical sociological theory to ecological issues is partly a manifestation of the way classical sociology was appropriated in the late 20th century. This can be viewed as *the appropriation problem.* Second, environmental sociology's critique of classical traditions has itself often been rooted in an overly restrictive conception of what constitutes environmental theorizing, reducing it to a narrow "dark green" perspective (as exemplified by the deep ecology tradition).[3] This can be thought of as *the definitional problem.*

## THE DEBATE ON MARX AND THE ENVIRONMENT

It is a sign of the growing influence of environmental issues that in recent years numerous thinkers, from Plato to Gandhi, have had their work reevaluated in relation to ecological analysis. Yet it is in relation to Marx's work that the largest and most controversial body of literature can be found, far overshadowing the debate over all other thinkers. This literature (insofar as it takes environmental issues seriously) has fallen into four camps: (1) those who contend that Marx's thought was antiecological from

---

[3] Environmentalists sometimes use the terms "dark green" and "light green" to refer to the same division as that between "deep ecology" and so-called "shallow ecology." In both cases, the nature of the distinction is the same: between what is thought of as an "anthropocentric" perspective versus a more "ecocentric" one—though such distinctions are notoriously difficult to define. For a sympathetic account of deep ecology, see McLaughlin (1993).

beginning to end and indistinguishable from Soviet practice (Clark 1989; Ferkiss 1993); (2) those who claim that Marx provided illuminating insights into ecology but ultimately succumbed to "Prometheanism" (protechnological, antiecological views)—a corollary being that he believed that environmental problems would be eliminated as a result of the "abundance" that would characterize postcapitalist society (Giddens 1981; Nove 1987; Redclift 1984; Benton 1989; McLaughlin 1990; Eckersley 1992; Deléage 1994; Goldblatt 1996); (3) those who argue that Marx provided an analysis of ecological degradation within agriculture, which remained, however, segregated off from his core social analysis (O'Connor 1998); and (4) those who insist that Marx developed a systematic approach to nature and to environmental degradation (particularly in relation to the fertility of the soil) that was intricately bound to the rest of his thought and raised the question of ecological sustainability (Parsons 1977; Perelman 1993; Mayumi 1991; Lebowitz 1992; Altvater 1993; Foster 1997; Burkett 1997).

Some of the sharpest criticisms of Marx from an environmentalist standpoint have come from leading sociologists (both non-Marxist and Marxist), particularly in Britain. Giddens (1981, p. 60) has contended that Marx, although demonstrating considerable ecological sensitivity in his earliest writings, later adopted a "Promethean attitude" toward nature. Marx's "concern with transforming the exploitative human social relations expressed in class systems does not extend," Giddens writes, "to the exploitation of nature" (1981, p. 59). Similarly, Redclift (1984, p. 7) has observed that for Marx the environment served "an enabling function but all value was derived from labor power. It was impossible to conceive of a 'natural' limit to the material productive forces of society. The barriers that existed to the full realization of resource potential were imposed by property relations and legal obligations rather than resource endowments." More recently, Redclift and Woodgate (1994, p. 53) have added that, "while Marx considered our relations with the environment as essentially social, he also regarded them as ubiquitous and unchanging, common to each phase of social existence. Hence, for Marx, the relationship between people and nature cannot provide a source of change in society. . . . Such a perspective does not fully acknowledge the role of technology, and its effects on the environment." Finally, Nove (1987, p. 399) has contended that Marx believed that "the problem of production had been 'solved'" by capitalism and that the future society of associated producers therefore would not have "to take seriously the problem of the allocation of scarce resources," which meant that there was no need for an "ecologically conscious" socialism.

Marx thus stands accused of wearing *blinders* in relation to the following: (1) the exploitation of nature, (2) nature's role in the creation of value,

(3) the existence of distinct natural limits, (4) nature's changing character and the impact of this on human society, (5) the role of technology in environmental degradation, and (6) the inability of mere economic abundance to solve environmental problems. If these criticisms were valid, Marx's work could be expected to offer no significant insights into problems of ecological crisis and indeed would itself constitute a major obstacle to the understanding of environmental problems.

In contrast, an attempt will be made to demonstrate here, in the context of a systematic reconstruction of Marx's theory of metabolic rift, that these ecological blinders are not in fact present in Marx's thought—and that each of the problems listed above were addressed to some extent in his theory. Of more significance, it will be contended that Marx provided a powerful analysis of the main ecological crisis of his day—the problem of soil fertility within capitalist agriculture—as well as commenting on the other major ecological crises of his time (the loss of forests, the pollution of the cities, and the Malthusian specter of overpopulation). In doing so, he raised fundamental issues about the antagonism of town and country, the necessity of ecological sustainability, and what he called the "metabolic" relation between human beings and nature. In his theory of metabolic rift and his response to Darwinian evolutionary theory, Marx went a considerable way toward a historical-environmental-materialism that took into account the coevolution of nature and human society.

## MARX AND THE SECOND AGRICULTURAL REVOLUTION: THE METABOLIC RIFT

### The Concept of the Second Agricultural Revolution

Although it is still common for historians to refer to a single agricultural revolution that took place in Britain in the 17th and 18th centuries and that laid the foundation for the industrial revolution that followed, agricultural historians commonly refer to a second and even a third agricultural revolution. The first agricultural revolution was a gradual process occurring over several centuries, associated with the enclosures and the growing centrality of market relations; technical changes included improved techniques of crop rotation, manuring, drainage, and livestock management. In contrast, the second agricultural revolution (Thompson 1968) occurred over a shorter period (1830–80) and was characterized by the growth of a fertilizer industry and a revolution in soil chemistry, associated in particular with the work of the great German agricultural chemist Justus von Liebig.[4] The third agricultural revolution was to occur still

[4] Thompson (1968) designates the second agricultural revolution as occurring over the years 1815–80, that is, commencing with the agricultural crisis that immediately followed the Napoleonic Wars. I have narrowed the period down to 1830–80 here in

later, in the 20th century, and involved the replacement of animal traction with machine traction on the farm and the eventual concentration of animals in massive feedlots, together with the genetic alteration of plants (resulting in narrower monocultures) and the more intensive use of chemical inputs—such as fertilizers and pesticides.

Marx's critique of capitalist agriculture and his main contributions to ecological thought have to be understood in relation to the second agricultural revolution occurring in his time. For Marx, writing in *Capital* in the 1860s, there was a gulf separating the treatment of agricultural productivity and soil fertility in the work of classical economists like Malthus and Ricardo, and the understanding of these problems in his own day. In Marx's ([1863–65] 1981, pp. 915–16) words, "The actual causes of the exhaustion of the land . . . were unknown to any of the economists who wrote about differential rent, on account of the state of agricultural chemistry in their time."

The source of the differential fertility from which rent was derived was, in the work of Malthus and Ricardo in the opening decades of the 19th century, attributed almost entirely to the natural or absolute productivity of the soil—with agricultural improvement (or degradation) playing only a marginal role. As Ricardo (1951, p. 67) observed, rent could be defined as "that portion of the produce of the earth, which is paid to the landlord for the use of the original and indestructible powers of the soil." These thinkers argued—with the presumed backing of natural law—that lands that were naturally the most fertile were the first to be brought into production and that rising rent on these lands and decreasing agricultural productivity overall were the result of lands of more and more marginal fertility being brought into cultivation, in response to increasing demographic pressures. Further, while some agricultural improvement was possible, it was quite limited, since the increases in productivity to be derived from successive applications of capital and labor to any given plot of land were said to be of diminishing character, thereby helping to account for the slowdown in growth of productivity in agriculture. All of this pointed to the Malthusian dilemma of a tendency of population to outgrow food supply—a tendency only countered as a result of vice and misery that served to lower fecundity and increase mortality, as Malthus emphasized in his original essay on population, or through possible moral restraint, as he was to add in later editions of that work.

---

order to distinguish between the crisis that to some extent preceded the second agricultural revolution and the revolution proper, for which the turning point was the publication of Liebig's *Organic Chemistry* in 1840 followed by J. B. Lawes's building of the first factory for the production of synthetic fertilizer (superphosphates) a few years later.

Classical Marxism, in contrast, relied from the beginning on the fact that rapid historical improvement in soil fertility was possible, though not inevitable, given existing social relations. In his "Outlines of a Critique of Political Economy," published in 1844, a young Friedrich Engels was to point to revolutions in science and particularly soil chemistry—singling out the discoveries of such figures as Humphry Davy and Liebig—as constituting the main reason why Malthus and Ricardo would be proven wrong about the possibilities for rapidly improving the fertility of the soil and thereby promoting a favorable relation between the growth of food and the growth of population. Engels (1964, pp. 208–10) went on to observe that, "To make earth an object of huckstering—the earth which is our one and all, the first condition of our existence—was the last step toward making oneself an object of huckstering." Three years later in *The Poverty of Philosophy,* Marx (1963, pp. 162–63) wrote that at "every moment the modern application of chemistry is changing the nature of the soil, and geological knowledge is just now, in our days, beginning to revolutionize all the old estimates of relative fertility. . . . Fertility is not so natural a quality as might be thought; it is closely bound up with the social relations of the time."

This emphasis on historical changes in soil fertility in the direction of agricultural improvement was to be a continuing theme in Marx's thought, though it eventually came to be coupled with an understanding of how capitalist agriculture could undermine the conditions of soil fertility, resulting in soil degradation rather than improvement. Thus in his later writings, increasing emphasis came to be placed on the exploitation of the earth in the sense of the failure to sustain the conditions of its reproduction.

## Liebig and the Depletion of the Soil

During 1830–70 the depletion of soil fertility through the loss of soil nutrients was the overriding environmental concern of capitalist society in both Europe and North America, comparable only to concerns over the growing pollution of the cities, deforestation of whole continents, and the Malthusian fears of overpopulation (Foster 1997; O'Connor 1998, p. 3). In the 1820s and 1830s in Britain, and shortly afterward in the other developing capitalist economies of Europe and North America, widespread concerns about "soil exhaustion" led to a phenomenal increase in the demand for fertilizer. The value of bone imports to Britain increased from £14,400 in 1823 to £254,600 in 1837. The first boat carrying Peruvian guano (accumulated dung of sea birds) unloaded its cargo in Liverpool in 1835; by 1841, 1,700 tons were imported, and by 1847, 220,000 (Ernle [1912] 1961, p. 369). European farmers in this period raided Napoleonic battlefields such

as Waterloo and Austerlitz, so desperate were they for bones to spread over their fields (Hillel 1991, pp. 131–32).

The second agricultural revolution associated with the rise of modern soil science was closely correlated with this demand for increased soil fertility to support capitalist agriculture. In 1837, the British Association for the Advancement of Science commissioned Liebig to write a work on the relationship between agriculture and chemistry. The following year saw the founding of the Royal Agricultural Society of England, viewed by economic historians as a leading organization in the British high-farming movement—a movement of wealthy landowners to improve farm management. In 1840, Liebig published his *Organic Chemistry in Its Applications to Agriculture and Physiology,* which provided the first convincing explanation of the role of soil nutrients, such as nitrogen, phosphorous, and potassium, in the growth of plants. One of the figures most influenced by Liebig's ideas was the wealthy English landowner and agronomist J. B. Lawes. In 1842, Lawes invented a means of making phosphate soluble, enabling him to introduce the first artificial fertilizer, and in 1843, he built a factory for the production of his new "superphosphates." With the repeal of the Corn Laws in 1846, Liebig's organic chemistry was seen by the large agricultural interests in England as the key to obtaining larger crop yields (Brock 1997, pp. 149–50).

In the 1840s, this scientific revolution in soil chemistry, together with the rise of a fertilizer industry, promised to generate a faster rate of agricultural improvement—impressing many contemporary observers, including Marx and Engels, who up to the 1860s believed that progress in agriculture might soon outpace the development of industry in general. Still, capital's ability to take advantage of these scientific breakthroughs in soil chemistry was limited by development of the division of labor inherent to the system, specifically the growing antagonism between town and country. By the 1860s, when he wrote *Capital,* Marx had become convinced of the contradictory and unsustainable nature of capitalist agriculture, due to two historical developments in his time: (1) the widening sense of crisis in agriculture in both Europe and North America associated with the depletion of the natural fertility of the soil, which was in no way alleviated, but rather given added impetus by the breakthroughs in soil science; and (2) a shift in Liebig's own work in the late 1850s and early 1860s toward an ecological critique of capitalist development.

The discoveries by Liebig and other soil scientists, while holding out hope to farmers, also intensified in some ways the sense of crisis within capitalist agriculture, making farmers more acutely aware of the depletion of soil minerals and the paucity of fertilizers. The contradiction was experienced with particular severity in the United States—especially among farmers in New York and in the plantation economy of the Southeast.

Blocked from ready access to guano (which was high in both nitrogen and phosphates) by the British monopoly of Peruvian guano supplies, U.S. capitalists spread across the globe looking for alternative supplies. Nevertheless, the quantity and quality of natural fertilizer obtained in this way fell far short of U.S. needs (Skaggs 1994).

Peruvian guano was largely exhausted in the 1860s and had to be replaced by Chilean nitrates. Potassium salts discovered in Europe gave ample access to that mineral, and phosphates became more readily available through both natural and artificial supplies. Yet prior to the development of a process for producing synthetic nitrogen fertilizer in 1913, fertilizer nitrogen continued to be in chronically short supply. It was in this context that Liebig was to state that what was needed to overcome this barrier was the discovery of "deposits of manure or guano . . . in volumes approximating to those of the English coalfields" (quoted in Kautsky [(1899) 1988], vol. 1, p. 53).

The second agricultural revolution, associated with the application of scientific chemistry to agriculture, was therefore at the same time a period of intense contradictions. The decline in the natural fertility of the soil due to the disruption of the soil nutrient cycle, the expanding scientific knowledge of the need for specific soil nutrients, and the simultaneous limitations in the supply of both natural and synthetic fertilizers, all served to generate serious concerns about present and future soil fertility under capitalist agriculture.

In upstate New York, increased competition from farmers to the west in the decades following the opening of the Erie Canal in 1825 intensified the concern over the "worn-out soil." In 1850, the British soil chemist, James F. W. Johnston, whom Marx (Marx and Engels 1975a, vol. 38, p. 476) was to call "the English Liebig," visited the United States. In his *Notes on North America,* Johnston (1851, pp. 356–65) recorded the depleted condition of the soil in upstate New York, comparing it unfavorably to the more fertile, less exhausted farmlands to the west. These issues were taken up by the U.S. economist Henry Carey, who in the late 1840s and 1850s laid stress on the fact that long-distance trade, which he associated with the separation of town from country and of agricultural producers from consumers ([1847] 1967a, pp. 298–99, 304–8), was the major factor in the net loss of nutrients to the soil and in the growing soil fertility crisis. "As the whole energies of the country," Carey wrote of the United States in his *Principles of Social Science,* "are given to the enlargement of the trader's power, it is no matter of surprise that its people are everywhere seen employed in 'robbing the earth of its capital stock'" ([1858–59] 1867, p. 215; also Carey [1853] 1967b, p. 199).

Carey's views were to have an important impact on Liebig. In his *Letters on Modern Agriculture* (1859), Liebig argued that the "empirical agri-

culture" of the trader gave rise to a "spoliation system" in which the "conditions of reproduction" of the soil were undermined. "A field from which something is permanently taken away," he wrote, "cannot possibly increase or even continue equal in its productive power." Indeed, "every system of farming based on the spoliation of the land leads to poverty" (1859, pp. 175–78). "Rational agriculture, in contrast to the spoliation system of farming, is based on the principle of restitution; by giving back to the fields the conditions of their fertility, the farmer insures the permanence of the latter." For Liebig, English "high farming" was "not the open system of robbery of the American farmer . . . but is a more refined species of spoliation which at first glance does not look like robbery" (1859, p. 183). Echoing Carey (1858), Liebig (1859, p. 220) observed that there were hundreds, sometimes thousands, of miles in the United States between the centers of grain production and their markets. The constituent elements of the soil were thus shipped to locations far removed from their points of origin, making the reproduction of soil fertility that much more difficult.

The problem of the pollution of the cities with human and animal wastes was also tied to the depletion of the soil. In Liebig's (1863, p. 261) words, "If it were practicable to collect, with the least loss, all the solid and fluid excrements of the inhabitants of the town, and return to each farmer the portion arising from produce originally supplied by him to the town, the productiveness of the land might be maintained almost unimpaired for ages to come, and the existing store of mineral elements in every fertile field would be amply sufficient for the wants of increasing populations." In his influential *Letters on the Subject of the Utilization of the Municipal Sewage* (1865) Liebig argued—basing his analysis on the condition of the Thames—that organic recycling that would return the nutrients contained in sewage to the soil was an indispensable part of a rational urban-agricultural system.

Marx and the Metabolic Rift

When working on *Capital* in the early 1860s, Marx was deeply affected by Liebig's analysis. In 1866, he wrote to Engels that in developing his critique of capitalist ground rent, "I had to plough through the new agricultural chemistry in Germany, in particular Liebig and Schönbein, which is more important for this matter than all the economists put together" (Marx and Engels 1975a, vol. 42, p. 227). Indeed, "to have developed from the point of view of natural science the negative, i.e., destructive side of modern agriculture," Marx was to note in *Capital*, "is one of Liebig's immortal merits" ([1867] 1976, p. 638). Far from having ecological blinders with regard to the exploitation of the earth, Marx, under the influence of Liebig's work of the late 1850s and early 1860s, was to develop

a systematic critique of capitalist "exploitation" (in the sense of robbery, i.e., failing to maintain the means of reproduction) of the soil.

Marx concluded both of his two main discussions of capitalist agriculture with an explanation of how large-scale industry and large-scale agriculture combined to impoverish the soil and the worker. Much of the resulting critique was distilled in a remarkable passage at the end of Marx's treatment of "The Genesis of Capitalist Ground Rent" in *Capital*, volume 3, where he wrote:

> Large landed property reduces the agricultural population to an ever decreasing minimum and confronts it with an ever growing industrial population crammed together in large towns; in this way it produces conditions that provoke an irreparable rift in the interdependent process of the social metabolism, a metabolism prescribed by the natural laws of life itself. The result of this is a squandering of the vitality of the soil, which is carried by trade far beyond the bounds of a single country. (Liebig.) . . . Large-scale industry and industrially pursued large-scale agriculture have the same effect. If they are originally distinguished by the fact that the former lays waste and ruins the labour-power and thus the natural power of man, whereas the latter does the same to the natural power of the soil, they link up in the later course of development, since the industrial system applied to agriculture also enervates the workers there, while industry and trade for their part provide agriculture with the means of exhausting the soil. (Marx 1981, pp. 949–50)

Marx provided a similar and no less important distillation of his critique in this area in his discussion of "Large-scale Industry and Agriculture" in volume 1 of *Capital:*

> Capitalist production collects the population together in great centres, and causes the urban population to achieve an ever-growing preponderance. This has two results. On the one hand it concentrates the historical motive force of society; on the other hand, it disturbs the metabolic interaction between man and the earth, i.e. it prevents the return to the soil of its constituent elements consumed by man in the form of food and clothing; hence it hinders the operation of the eternal natural condition for the lasting fertility of the soil. . . . But by destroying the circumstances surrounding that metabolism . . . it compels its systematic restoration as a regulative law of social production, and in a form adequate to the full development of the human race. . . . All progress in capitalist agriculture is a progress in the art, not only of robbing the worker, but of robbing the soil; all progress in increasing the fertility of the soil for a given time is a progress toward ruining the more long-lasting sources of that fertility. . . . Capitalist production, therefore, only develops the techniques and the degree of combination of the social process of production by simultaneously undermining the original sources of all wealth—the soil and the worker. (Marx 1976, pp. 637–38)

In both of these passages from Marx's *Capital*—the first concluding his discussion of capitalist ground rent in volume 3 and the second concluding

his discussion of large-scale agriculture in volume 1—the central theoretical construct is that of a "rift" in the "metabolic interaction between man and the earth," or in the "social metabolism prescribed by the natural laws of life," through the removal from the soil of its constituent elements, requiring its "systematic restoration." This contradiction is associated with the growth simultaneously of large-scale industry and large-scale agriculture under capitalism, with the former providing agriculture with the means of the intensive exploitation of the soil. Following Liebig, Marx argued that long-distance trade in food and clothing made the problem of the alienation of the constituent elements of the soil that much more of an "irreparable rift." As he indicated elsewhere in *Capital* (vol. 1), the fact that "the blind desire for profit" had "exhausted the soil" of England could be seen daily in the conditions that "forced the manuring of English fields with guano" imported from Peru (1976, p. 348). Central to Marx's argument was the notion that capitalist large-scale agriculture prevents any truly rational application of the new science of soil management. Despite all of its scientific and technological development in the area of agriculture, capitalism was unable to maintain those conditions necessary for the recycling of the constituent elements of the soil.

The key to Marx's entire theoretical approach in this area is the concept of social-ecological metabolism (*Stoffwechsel*), which was rooted in his understanding of the labor process. Defining the labor process in general (as opposed to its historically specific manifestations), Marx employed the concept of metabolism to describe the human relation to nature through labor:

> Labour is, first of all, a process between man and nature, a process by which man, through his own actions, mediates, regulates and controls the metabolism between himself and nature. He confronts the materials of nature as a force of nature. He sets in motion the natural forces which belong to his own body, his arms, legs, head and hands, in order to appropriate the materials of nature in a form adapted to his own needs. Through this movement he acts upon external nature and changes it, and in this way he simultaneously changes his own nature. . . . It [the labor process] is the universal condition for the metabolic interaction [*Stoffwechsel*] between man and nature, the everlasting nature-imposed condition of human existence. (Marx 1976, pp. 283, 290)

Only a few years before this, Marx had written in his *Economic Manuscript of 1861–63* that "actual labour is the appropriation of nature for the satisfaction of human needs, the activity through which the metabolism between man and nature is mediated." It followed that the actual activity of labor was never independent of nature's own wealth-creating potential, "since material wealth, the world of use values, exclusively consists of

natural materials modified by labour" (Marx and Engels 1975a, vol. 30, p. 40).[5]

Much of this discussion of the metabolic relation between human beings and nature reflected Marx's early, more directly philosophical attempts to account for the complex interdependence between human beings and nature. In the *Economic and Philosophical Manuscripts* of 1844, Marx had explained that, "Man *lives* from nature, i.e., nature is his *body*, and he must maintain a continuing dialogue with it if he is not to die. To say that man's physical and mental life is linked to nature simply means that nature is linked to itself, for man is a part of nature" (1974, p. 328; emphasis in original). But the later introduction of the concept of metabolism gave Marx a more solid—and scientific—way in which to depict the complex, dynamic interchange between human beings and nature, resulting from human labor. The material exchanges and regulatory action associated with the concept of metabolism encompassed both "nature-imposed conditions" and the capacity of human beings to affect this process. According to Hayward (1994, p. 116), Marx's concept of socio-ecological metabolism "captures fundamental aspects of humans' existence as both natural and physical beings: these include the energetic and material exchanges which occur between human beings and their natural environment. . . . This metabolism is regulated from the side of nature by natural laws governing the various physical processes involved, and from the side of society by institutionalized norms governing the division of labor and distribution of wealth etc."

Given the fundamental way in which Marx conceived of the concept of metabolism—as constituting the complex, interdependent process linking human society to nature—it should not surprise us that this concept enters

[5] Marx highlighted the methodological importance of the concept of "material exchange [*Stoffwechsel*] between man and nature" in his *Notes on Adolph Wagner*, his last economic work, written in 1880 (1975, p. 209). As early as 1857–58 in the *Grundrisse*, Marx had referred to the concept of metabolism (*Stoffwechsel*) in the wider sense of "a system of general social metabolism, of universal relations, of all-round needs and universal capacities . . . formed for the first time" under generalized commodity production (1973, p 158). Throughout his later economic works, he employed the concept to refer both to the actual metabolic interaction between nature and society through human labor, and also in a wider sense to describe the complex, dynamic, interdependent set of needs and relations brought into being and constantly reproduced in alienated form under capitalism, and the question of human freedom that this raised—all of which could be seen as being connected to the way in which the human metabolism with nature was expressed through the organization of human labor. Marx thus gave the concept of metabolism both a specific ecological meaning and a wider social meaning. It makes sense therefore to speak of the "socioecological" nature of his concept.

into Marx's vision of a future society of associated producers: "Freedom, in this sphere [the realm of natural necessity]," he wrote in *Capital* (volume 3), "can consist only in this, that socialized man, the associated producers, govern the human metabolism with nature in a rational way, bringing it under their own collective control rather than being dominated by it as a blind power; accomplishing it with the least expenditure of energy and in conditions most worthy and appropriate for their human nature" (1981, p. 959).

Just as the introduction of the concept of "metabolism" allowed Marx to provide a firmer, scientific grounding for his ideas, so the central position that this concept came to occupy in his theory encouraged him to draw out some of its larger implications. The term "metabolism" (*Stoffwechsel*) was introduced as early as 1815 and was adopted by German physiologists in the 1830s and 1840s to refer to material exchanges within the body, related to respiration (Bing 1971; Caneva 1993). But the term was given a somewhat wider application (and therefore greater currency) in 1842 by Liebig in his *Animal Chemistry*, the great work that followed his earlier work on the soil, where he introduced the notion of metabolic process (in the context of tissue degradation). It was subsequently generalized still further and emerged as one of the key concepts, applicable both at the cellular level and in the analysis of entire organisms, in the development of biochemistry (Liebig [1842] 1964; Brock 1997, p. 193; Caneva 1993, p. 117).

Within biological and ecological analysis, the concept of metabolism, beginning in the 1840s and extending down to the present day, has been used as a central category in the systems-theory approach to the relation of organisms to their environments. It refers to a complex process of metabolic exchange, whereby an organism (or a given cell) draws upon materials and energy from its environment and converts these by way of various metabolic reactions into the building blocks of proteins and other compounds necessary for growth. The concept of metabolism is also used to refer to the regulatory processes that govern this complex interchange between organisms and their environment (Fischer-Kowalski 1997, p. 120). Leading system ecologists like Odum (1969, p. 7) employ "metabolism" to refer to all biological levels, beginning with the single cell and ending with the ecosystem.

Recently, the notion of metabolism has become what Fischer-Kowalski (1997, pp. 119–20) has called "a rising conceptual star" within social-ecological thought, as a result of the emergence of cross-disciplinary research in "industrial metabolism." For some thinkers, it offers a way out of one the core dilemmas of environmental sociology raised by Dunlap and Catton (1979) and Schnaiberg (1980), which requires a way of envisioning the complex interaction between society and nature (Hayward 1994, pp. 116–

17; Fischer-Kowalski 1997). Further, the concept of metabolism has long been employed to analyze the material interchange between city and country, in a manner similar to the way in which Liebig and Marx used the concept (Wolman 1965; Giradet 1997). Within this rapidly growing body of literature on social-ecological metabolism, it is now well recognized that "within the nineteenth-century foundations of social theory, it was Marx and Engels who applied the term 'metabolism' to society" (Fischer-Kowalski 1997, p. 122).

Indeed, environmental sociologists and others exploring the concept of "industrial metabolism" today argue that just as the materials that birds use to build their nests can be seen as material flows associated with the metabolism of birds, so similar material flows can be seen as part of the human metabolism. Fischer-Kowalski has thus suggested "considering as part of the metabolism of a social system *those material and energetic flows that sustain the material compartments of the system*" (1997, pp. 121, 131; emphasis in original). The tough question, however, is how such a human metabolism with nature is regulated on the side of society. For Marx, the answer was human labor and its development within historical social formations.

MARX AND SUSTAINABILITY

An essential aspect of the concept of metabolism is the notion that it constitutes the basis on which life is sustained and growth and reproduction become possible. Contrary to those who believe that he wore an ecological blinder that prevented him from perceiving natural limits to production, Marx employed the concept of metabolic rift to capture the material estrangement of human beings in capitalist society from the natural conditions of their existence. To argue that large-scale capitalist agriculture created such a metabolic rift between human beings and the soil was to argue that basic conditions of sustainability had been violated. "Capitalist production," Marx ([1861–63] 1971*b*, p. 301) wrote, "turns toward the land only after its influence has exhausted it and after it has devastated its natural qualities." Moreover, this could be seen as related not only to the soil but to the antagonism between town and country. For Marx, like Liebig, the failure to recycle nutrients to the soil had its counterpart in the pollution of the cities and the irrationality of modern sewage systems. In *Capital* (volume 3), he observed: "In London . . . they can do nothing better with the excrement produced by 4 1/2 million people than pollute the Thames with it, at monstrous expense" (1981, p. 195). Engels was no less explicit on this point. In addressing the need to transcend the antagonism between town and country, he referred, following Liebig, to the fact that "in London alone a greater quantity of manure than is produced by

the whole kingdom of Saxony is poured away every day into the sea with an expenditure of enormous sums" and to the consequent need to reestablish an "intimate connection between industrial and agricultural production" along with "as uniform a distribution as possible of the population over the whole country" (Engels [1872] 1975, p. 92). For Marx, the "excrement produced by man's natural metabolism," along with the waste of industrial production and consumption, needed to be recycled back into the production, as part of a complete metabolic cycle (1981, p. 195).

The antagonistic division between town and country, and the metabolic rift that it entailed, was also evident at a more global level: whole colonies saw their land, resources, *and soil* robbed to support the industrialization of the colonizing countries. "For a century and a half," Marx wrote, "England has indirectly exported the soil of Ireland, without as much as allowing its cultivators the means for making up the constituents of the soil that had been exhausted" (1976, p. 860).

Marx's view of capitalist agriculture and of the necessity of cycling the nutrients of the soil (including the organic wastes of the city) thus led him to a wider concept of ecological sustainability—a notion that he thought of very limited practical relevance to capitalist society, which was incapable of such consistent rational action, but essential for a future society of associated producers. "The way that the cultivation of particular crops depends on fluctuations in market prices and the constant change in cultivation with these prices—the entire spirit of capitalist production, which is oriented towards the most immediate monetary profits—stands in contradiction to agriculture, which has to concern itself with the whole gamut of permanent conditions of life required by the chain of successive generations" (Marx 1981, p. 754).

In emphasizing the need to maintain the earth for "successive generations," Marx captured the essence of the contemporary notion of sustainable development, defined most famously by the Brundtland Commission as "development which meets the needs of the present without compromising the ability of future generations to meet their needs" (World Commission on Environment and Development 1987, p. 43). For Marx, the "conscious and rational treatment of the land as permanent communal property" is "the inalienable condition for the existence and reproduction of the chain of human generations" (1981, pp. 948–49). Indeed, in a remarkable, and deservedly famous, passage in *Capital* (vol. 3), Marx wrote, "From the standpoint of a higher socio-economic formation, the private property of particular individuals in the earth will appear just as absurd as the private property of one man in other men. Even an entire society, a nation, or all simultaneously existing societies taken together, are not owners of the earth, they are simply its possessors, its beneficiaries, and

have to bequeath it in an improved state to succeeding generations as *boni patres familias* [good heads of the household]" (1981, p. 911).

This took on greater significance near the end of Marx's life, when, as a result of his investigations into the revolutionary potential of the archaic Russian commune (the Mir), he argued that it would be possible to develop an agricultural system "organized on a vast scale and managed by cooperative labor" through the introduction of "modern agronomic methods." The value of such a system, he argued, would be that it would be "in a position to incorporate all the positive acquisitions devised by the capitalist system" without falling prey to the purely exploitative relation to the soil, that is, the robbery, that characterized the latter (Marx and Engels 1975*a*, vol. 24, p. 356). Marx's absorption in the literature of the Russian populists at the end of his life, and his growing conviction that the revolution would emerge first within Russia—where economic, and more specifically agricultural, abundance could not be assumed—forced him to focus on agricultural underdevelopment and the ecological requirements of a more rational agricultural system.[6]

Marx and Engels did not restrict their discussions of environmental degradation to the robbing of the soil but also acknowledged other aspects of this problem, including the depletion of coal reserves, the destruction of forests, and so on. As Engels observed in a letter to Marx, "the working individual is not only a stabaliser of *present* but also, and to a far greater extent, a squanderer of *past,* solar heat. As to what we have done in the way of squandering our reserves of energy, our coal, ore, forests, etc., you are better informed than I am" (Marx and Engels 1975*a*, vol. 46, p. 411; emphasis in original). Marx referred to the "devastating" effects of "deforestation" (Marx and Engels 1975*a*, vol. 42, p. 559) and saw this as a long-term result of an exploitative relation to nature (not simply confined to capitalism): "The development of civilization and industry in general," Marx wrote, "has always shown itself so active in the destruction of forests that everything that has been done for their conservation and production is completely insignificant in comparison" ([1865–70] 1978, p. 322). He lamented the fact that the forests in England were not "true forests" since "the deer in the parks of the great are demure domestic cattle, as fat as London aldermen"; while in Scotland, the so-called "deer-forests" that were established for the benefit of huntsmen (at the expense of rural laborers) contained deer but no trees (1976, pp. 892–93). Under the influence of Darwin, Marx and Engels repudiated the age-old view that human beings were at the center of the natural universe. Engels expressed "a

[6] On this later phase of Marx's analysis, in which he addressed the agricultural concerns of the Russian populists, see Shanin (1983).

withering contempt for the idealistic exaltation of man over the other animals" (Marx and Engels 1975*b*, p. 102).

Some critics attribute to Marx an ecological blinder associated with an overly optimistic faith in the cornucopian conditions supposedly made possible by the forces of production under capitalism. In this view, he relied so much on the assumption of abundance in his conception of a future society that ecological factors such as the scarcity of natural resources were simply nonexistent. Yet whatever Marx may have thought in his more "utopian" conceptions, it is clear from his discussions of both capitalism and of the transition to socialism that he was far from believing, as Nove (1987, p. 399) contends, "that the problem of production" had already been "solved" under capitalism or that natural resources were "inexhaustible." Rather, capitalism, as he emphasized again and again, was beset with a chronic problem of production in agriculture, which ultimately had to do with an unsustainable form of production in relation to natural conditions. Agriculture, Marx observed, "when it progresses spontaneously and is not *consciously controlled* . . . leaves deserts behind it" (Marx and Engels 1975*b*, p. 190; emphasis in original). Within industry too, Marx was concerned about the enormous waste generated and emphasized the "reduction" and "re-use" of waste—particularly in a section of *Capital* (volume 3), entitled, "Utilization of the Refuse of Production" (1981, pp. 195–97). Moreover, he gave every indication that these problems would continue to beset any society attempting to construct socialism (or communism). Hence, although some critics, such as McLaughlin (1990, p. 95), assert that Marx envisioned "a general material abundance as the substratum of communism," and therefore saw "no basis for recognizing any interest in the liberation of nature from human domination," overwhelming evidence to the contrary (much of it referred to above) suggests that Marx was deeply concerned with issues of ecological limits and sustainability.

Moreover, there is simply no indication anywhere in Marx's writings that he believed that a sustainable relation to the earth would come automatically with the transition to socialism. Rather, he emphasized the need for planning in this area, including such measures as the elimination of the antagonism between town and country through the more even dispersal of the population (Marx and Engels [1848] 1967, pp. 40–41) and the restoration and improvement of the soil through the recycling of soil nutrients. All of this demanded a radical transformation in the human relation to the earth via changed production relations. Capitalism, Marx wrote, "creates the material conditions for a new and higher synthesis, a union of agriculture and industry on the basis of the forms that have developed during the period of their antagonistic isolation" (1976, p. 637). But in order to achieve this "higher synthesis" in a society of freely associated

producers, he argued, it would be necessary for the associated producers to "govern the human metabolism with nature in a rational way"—a requirement that raised fundamental challenges for postcapitalist society (1981, p. 959; 1976, pp. 637–38).

Another ecological blinder commonly attributed to Marx is that he denied the role of nature in the creation of wealth by developing a labor theory of value that saw all value as derived from labor, and by referring to nature as a "free gift" to capital, lacking any intrinsic value of its own (Deléage 1994, p. 48; Churchill 1996, pp. 467–68; Georgescu-Roegen 1971, p. 2). Yet this criticism is based on a misunderstanding of Marx's political economy. Marx did not invent the idea that the earth was a "gift" of nature to capital. This notion was advanced as a key proposition by Malthus and Ricardo in their economic works (Malthus 1970, p. 185). It was taken up later on by the great neoclassical economist Alfred Marshall (1920) and persisted in neoclassical economics textbooks into the 1980s. Thus, in the 10th edition of a widely used introductory economics textbook, we discover the following: "Land refers to all natural resources—all 'free gifts of nature'—which are usable in the production process." And further on we read, "Land has no production cost; it is a 'free and nonreproducible gift of nature' " (McConnell 1987, pp. 20, 672). Marx was aware of the social-ecological contradictions embedded in such views, and in his *Economic Manuscript of 1861–63* he attacked Malthus repeatedly for falling back on the "physiocratic" notion that the environment was "a gift of nature to man," while ignoring how this was connected to the definite set of social relations brought into being by capital (Marx and Engels 1975a, vol. 34, pp. 151–59).

To be sure, Marx agreed with liberal economics that under the law of value of capitalism nature was accorded no value. "The earth . . . is active as agent of production in the production of a use-value, a material product, say wheat," he wrote. "But it has nothing to do with producing the *value of the wheat*" (1981, p. 955). The value of the wheat as in the case of any commodity under capitalism was derived from labor. For Marx, however, this merely reflected the narrow, limited conception of wealth embodied in capitalist commodity relations and in a system built around exchange value. Genuine wealth consisted of use values—the characteristic of production in general, transcending its capitalist form. Hence, nature, which contributed to the production of use values, was just as much a source of wealth as labor. "What Lucretius says," Marx wrote in *Capital* (1976, p. 323), "is self-evident: *nil posse creari de nihilo,* out of nothing, nothing can be created. . . . Labour-power itself is, above all else, the material of nature transposed into a human organism."

It follows that "labour," as Marx stated at the beginning of *Capital,* "is not the only source of material wealth, that is, of the use-values it pro-

duces. As William Petty says, labour is the father of material wealth, and the earth is its mother" (1976, p. 134). In the *Critique of the Gotha Programme,* Marx criticized those socialists who had attributed what he called *"supernatural creative power* to labour" ([1875] 1971*a*, p. 11; emphasis in original) by viewing it as the sole source of wealth and disregarding the role of nature. Under communism, he argued, wealth would need to be conceived in far more universal terms, as consisting of those material use values that constituted the basis for the full development of human creative powers, "the development of the rich individuality which is all sided in its production as in its consumption"—expanding the wealth of connections allowed for by nature, while at the same time reflecting the developing human metabolism with nature ([1857–58] 1973, p. 325).

Marx therefore set himself in opposition to all those who thought the contribution of nature to the production of wealth could be disregarded, or that nature could be completely subordinated to human ends regardless of their character. Commenting in the *Grundrisse* on Bacon's ([1620] 1994, pp. 29, 43) great maxim that "nature is only overcome by obeying her"— on the basis of which Bacon also proposed to "subjugate" nature—Marx replied that for capitalism the theoretical discovery of nature's "autonomous laws appears merely as a ruse so as to subjugate it under human needs, whether as an object of consumption or a means of production" (1973, pp. 409–10).

For Engels too, it was clear that to construct a society built on the vain hope of the total conquest of external nature was sheer folly. As he wrote in *The Dialectics of Nature* ([1874–80] 1940, pp. 291–92), "Let us not, however, flatter ourselves overmuch on account of our human conquest of nature. For each such conquest takes revenge on us. . . . At every step we are reminded that we by no means rule over nature like a conqueror over a foreign people, like someone standing outside nature—but that we, with flesh, blood, and brain, belong to nature, and exist in its midst, and that all our mastery of it consists in the fact that we have the advantage of all other beings of being able to know and correctly apply its laws."

For Marx, "the human metabolism with nature" was a highly dynamic relationship, reflecting changes in the ways human beings mediated between nature and society through production. Engels and Marx read *The Origin of Species* soon after it appeared in 1859 and were enthusiastic supporters of Darwin's theory of natural selection. Marx (1976, p. 461) called Darwin's book an "epoch-making work," and in January 1861, Marx wrote a letter to the German socialist Ferdinand Lasalle stating that Darwin had dealt the "death blow" to " 'teleology' in the natural sciences" (Marx and Engels 1975*a*, vol. 41, pp. 246–47). Marx expressed no reservations about Darwin's fundamental theory itself—not even with regard to

Darwin's application of the Malthusian "struggle for existence" to the world of plants and animals—yet he was sharply critical of all attempts by social Darwinists to carry this analysis beyond its proper domain and to apply it to human history. Unfortunately, some critics have viewed his cautionary notes in this respect as criticisms of Darwin himself.[7]

Darwin's evolutionary theory led Marx and Engels to what would now be called a "cautious constructionism" (Dunlap 1997, pp. 31–32). For Marx, human evolution, that is, human history, was distinct from evolution as it occurred among plants and animals, in that the natural evolution of the physical organs of the latter, that is, "the history of natural technology," had its counterpart in human history in the conscious development of the "productive organs of man in society" (technology), which helped establish the conditions for the human mediation between nature and society via production (Marx 1976, p. 493). Marx was of course aware that the Greek word organ (organon) also meant tool, and that organs were initially viewed as "grown-on" tools of animals—an approach that was utilized by Darwin himself, who compared the development of specialized organs to the development of specialized tools (see Pannekoek 1912; Darwin [1859] 1968, pp. 187–88).

Engels was later to add to this an analysis of "The Part Played by Labour in the Transition from Ape to Man" (Engels 1940, pp. 279–96). According to this theory (verified in the 20th century by the discovery of *Australopithecus*), erect posture developed first (prior to the evolution of the human brain), freeing the hands for tools. In this way, the human (hominid) relation to the local environment was radically changed, altering the basis of natural selection. Those hominids that were most successful at toolmaking were best able to adapt, which meant that the evolutionary process exerted selective pressures toward the development of the brain, eventually leading to the rise of modern humans. The human brain, according to Engels, evolved then through a complex, interactive process, now referred to as "gene-culture evolution." As biologist and paleontolo-

---

[7] Marx and Engels's complex relation to Darwin's work—which neither denied a relation between society and biology nor reduced one to the other—may also have something to say about why they never utilized the term "ecology," coined by Darwin's leading German follower Ernst Haeckel in 1866, the year before the publication of volume 1 of *Capital*. Although the concept of ecology only gradually came into common usage, Marx and Engels were very familiar with Haeckel's work and so may have been aware of his coinage of this concept. Yet, the way that Haeckel, a strong social Darwinist, originally defined the term was unlikely to have predisposed them to its acceptance. "By ecology," Haeckel had written, "we mean the body of knowledge concerning the economy of nature . . in a word, ecology is the study of all those complex interrelations referred to by Darwin as the conditions of the struggle for existence" (Golley 1993, p. 207).

gist Stephen Jay Gould has observed, all scientific explanations of the evolution of the human brain thus far have taken the form of gene-culture coevolution, and "the best nineteenth-century case for gene-culture coevolution was made by Friedrich Engels" (Gould 1987, pp. 111–12). The analysis of Marx and Engels thus pointed to coevolution (Norgaard 1994), neither reducing society to nature, nor nature to society, but exploring their interactions. Indeed, the view that "nature reacts on man and natural conditions everywhere exclusively determined his historical development," Engels observed, "is . . . one-sided and forgets that man also reacts on nature, changing it and creating new conditions of existence for himself" (1940, p. 172).

The key to the metabolic relation of human beings to nature then is technology, but technology as conditioned by both social relations and natural conditions. Contrary to those who argue that Marx wore an ecological blinder when it came to envisioning the limitations of technology in surmounting ecological problems, he explicitly argued in his critique of capitalist agriculture, that while capitalism served to promote "technical development in agriculture," it also brought into being social relations that were "incompatible" with a sustainable agriculture (1981, p. 216). The solution thus lay less in the application of a given technology than in the transformation of social relations. Moreover, even if the most advanced technical means available were in the hands of the associated producers, nature, for Marx, sets certain limits. The reproduction of "plant and animal products," for example, is conditioned by "certain organic laws involving naturally determined periods of time" (1981, p. 213). Marx reiterated the Italian political economist Pietro Verri's statement that human production was not properly an act of creation but merely "the reordering of matter" and was thus dependent on what the earth provided (1976, p. 132). The human interaction with nature always had to take the form of a metabolic cycle that needed to be sustained for the sake of successive generations. Technological improvements were a necessary but insufficient means for the "improvement" in the human relation to the earth. For Marx, human beings transformed their relation to nature but not exactly as they pleased; they did so in accordance with conditions inherited from the past and as a result of a complex process of historical development that reflected a changing relation to a natural world, which was itself dynamic in character. Redclift and Woodgate (1994, p. 53) are therefore wrong when they say that Marx wore blinders in relation to the coevolution of nature and society, viewing the human relation to nature as an "unchanging" one. Engels began his *Dialectics of Nature* with a dramatic description of the historic defeat of 18th-century conceptions of nature in which the natural world existed only in space not in time; "in which all change, all development of nature was denied" (1940, p. 6).

## BEYOND THE APPROPRIATION AND DEFINITIONAL PROBLEMS

The foregoing suggests that Marx's analysis provides a multilayered and multivalent basis for linking sociology (and in particular the classical tradition of sociology) with environmental issues. Yet, if this is so, why has this concern with ecological issues not found a strong echo in the Marxist tradition throughout its development, and why has our understanding of Marx so often excluded these issues? Why has environmental sociology, which is concerned directly with these questions, been so slow to acknowledge Marx's importance in this respect? The first question relates to what we referred to at the beginning of this article as "the appropriation problem," the second to what was labeled "the definitional problem."

### The Appropriation Problem

Marx's reputation as an ecological thinker was no doubt affected by the fact that, as Massimo Quaini (1982, p. 136) has pointed out, he "denounced the spoliation of nature before a modern bourgeois ecological conscience was born." Nevertheless, Marx's ecological critique was fairly well-known and had a direct impact on Marxism in the decades immediately following his death. It came to be discarded only later on, particularly within Soviet ideology, as the expansion of production at virtually any cost became the overriding goal of the Communist movement. The influence of Marx's critique in this respect can be seen in the writings of such leading Marxist thinkers as Kautsky, Lenin, and Bukharin.

Kautsky's great work, *The Agrarian Question,* published in 1899, contained a section on "The Exploitation of the Countryside by the Town" in which he held that the net external flow of value from countryside to town "corresponds to a constantly mounting loss of nutrients in the form of corn, meat, milk and so forth which the farmer has to sell to pay taxes, debt-interest and rent. . . . Although such a flow does not signify an exploitation of agriculture in terms of the law of value [of the capitalist economy], it does nevertheless lead . . . to its material exploitation, to the impoverishment of the land of its nutrients" (Kautsky 1988 [1899], p. 214).[8] Arguing at a time when the fertilizer industry was further developed than

---

[8] In saying there was no exploitation of agriculture in law of value terms, Kautsky was arguing that transactions here, as in other areas of the economy, were based on equal exchange. Nonetheless, he insisted that "material exploitation" (related to use values) was present insofar as the soil was being impoverished. Marx too argued that the soil was being "robbed" or "exploited" in the latter sense and connected this to the fact that the land under capitalism was regarded as a "free gift" (as Malthus had contended) so that the full costs of its reproduction never entered into the law of value under capitalism.

in Marx's day, Kautsky discussed the fertilizer treadmill resulting from the metabolic rift:

> Supplementary fertilisers . . allow the reduction in soil fertility to be avoided, but the necessity of using them in larger and larger amounts simply adds a further burden to agriculture—not one unavoidably imposed by nature, but a direct result of current social organization. By overcoming the antithesis between town and country . . . the materials removed from the soil would be able to flow back in full. Supplementary fertilisers would then, at most, have the task of enriching the soil, not staving off its impoverishment. Advances in cultivation would signify an increase in the amount of soluble nutrients in the soil without the need to add artificial fertilisers. (Kautsky 1988, vol. 2, pp. 214–15)

Some of the same concerns were evident in Lenin's work. In *The Agrarian Question and the "Critics of Marx,"* written in 1901, he observed that, "The possibility of substituting artificial for natural manures and the fact that this is already being done (partly) do not in the least refute the irrationality of wasting natural fertilisers and thereby polluting the rivers and the air in suburban factory districts. Even at the present time there are sewage farms in the vicinity of large cities which utilise city refuse with enormous benefit to agriculture; but by this system only an infinitesimal part of the refuse is utilized" (1961, pp. 155–56).

It was Bukharin, however, who developed the most systematic approach to ecological issues in his chapter on "The Equilibrium between Society and Nature" in *Historical Materialism* his important work of the 1920s. Cohen (1980, p. 118) has characterized Bukharin's position as one of " 'naturalistic' materialism," because of its emphasis on the interaction between society and nature. As Bukharin wrote,

> This material process of "metabolism" between society and nature is the fundamental relation between environment and system, between "external conditions" and human society. . . . The metabolism between man and nature consists, as we have seen, in the transfer of material energy from external nature to society. . . . Thus, the interrelation between society and nature is a process of social reproduction. In this process, society applies its human labor energy and obtains a certain quantity of energy from nature ("nature's material," in the words of Marx). The *balance* between expenditures and receipts is here obviously the decisive element for the growth of society. If what is obtained exceeds the loss by labor, important consequences obviously follow for society, which vary with the amount of this excess. (Bukharin 1925, pp. 108–12)

For Bukharin, technology was the chief mediating force in this metabolic relationship between nature and society. The human metabolism with nature was thus an "unstable equilibrium," one which could be progressive or regressive from the standpoint of human society. "The productivity of labor," he wrote, "is a precise measure of the 'balance' between

society and nature." An increase in social productivity was seen as a progressive development; conversely, if the productivity of labor decreased—here Bukharin cited "the exhaustion of the soil" as a possible cause of such a decline—the relationship was a regressive one. Such a decline in social productivity resulting from an ill-adapted metabolic relation between society and nature could, he argued, lead to society being "barbarianized" (1925, pp. 77, 111–13).

Thus the whole "process of social production," Bukharin (1925, p. 111) wrote, "is an adaptation of human society to external nature." "Nothing could be more incorrect than to regard nature from the teleological point of view: man, the lord of creation, with nature created for his use, and all things adapted to human needs" (1925, p. 104). Instead, human beings were engaged in a constant, active struggle to adapt. "Man, as an animal form, as well as human society, are products of nature, part of this great, endless whole. Man can never escape from nature, and even when he 'controls' nature, he is merely making use of the laws of *nature* for his own ends" (1925, p. 104). "No system, including that of human society," Bukharin (1925, p. 89) insisted, "can exist in empty space; it is surrounded by an 'environment,' on which all its conditions ultimately depend. If human society is not adapted to its environment, it is not meant for this world." "For the tree in the forest, the environment means all the other trees, the brook, the earth, the ferns, the grass, the bushes, together with all their properties. Man's environment is society, in the midst of which he lives; the environment of human society is external nature" (1925, p. 75). Indeed, human beings, as Bukharin emphasized in 1931, need to be conceived as "living and working in the biosphere" (1971, p. 17).[9]

Other early Soviet thinkers connected to Bukharin demonstrated a similar concern for ecological issues. Komrov (1935, pp. 230–32) quoted at length from the long passage on the illusion of the conquest of nature in Engels's *Dialectics of Nature* and went on to observe that, "The private owner or employer, however necessary it may be to make the changing of the world comply with the laws of Nature, cannot do so since he aims at profit and only profit. By creating crisis upon crisis in industry he lays waste natural wealth in agriculture, leaving behind a barren soil and in mountain districts bare rocks and stony slopes." Similarly, Uranovsky

[9] In referring to the "biosphere," Bukharin drew upon V. I. Vernadsky's *The Biosphere*, first published in 1922, which was one of the great works in ecological science of the 20th century and was extremely influential in Soviet scientific circles in the 1920s and early 1930s. Vernadsky was "the first person in history to come [to] grips with the real implications of the fact that the Earth is a self-contained sphere" (Margulis et al 1998, p. 15). He achieved international renown both for his analysis of the biosphere and as the founder of the science of geochemistry (or biogeochemistry) (Vernadsky [1922] 1998).

(1935, p. 147) placed heavy emphasis, in a discussion of Marxism and science, on Marx's research into Liebig and "the theory of the exhaustion of the soil."[10]

Burkharin's ecological work and that of those associated with him was a product of the early Soviet era. The tragedy of the Soviet relation to the environment, which was eventually to take a form that has been characterized as "ecocide" (Feshbach and Friendly 1992; Peterson 1993), has tended to obscure the enormous dynamism of early Soviet ecology of the 1920s and the role that Lenin personally played in promoting conservation. In his writings and pronouncements, Lenin insisted that human labor could never substitute for the forces of nature and that a "rational exploitation" of the environment, or the scientific management of natural resources, was essential. As the principal leader of the young Soviet state, he argued for "preservation of the monuments of nature" and appointed the dedicated environmentalist Anatolii Vasil'evich Lunacharskii as head of the People's Commissariat of Education (Enlightenment), which was put in charge of conservation matters for all of Soviet Russia (Weiner 1988a, pp. 4, 22–28, 259; Weiner 1988b, pp. 254–55; Bailes 1990, pp. 151–58). Lenin had considerable respect for V. I. Vernadsky, the founder of the science of geochemistry (or biogeochemistry) and the author of *The Biosphere.* It was in response to the urging of Vernadsky and mineralogist E. A. Fersman that Lenin in 1919 established in the southern Urals the first nature preserve in the USSR—and indeed the first reserve anywhere by a government exclusively aimed at the scientific study of nature (Weiner 1988a, p. 29; Bailes 1990, p. 127). Under Lenin's protection, the Soviet conservation movement prospered, particularly during the New Economic Policy period (1921–28). But with the early death of Lenin and the triumph of Stalinism in the late 1920s, conservationists were attacked for being "bourgeois." Worse still, with the rise of Trofim Denisovich Lysenko, as an arbiter of biological science, "scientific" attacks were launched first on ecology and then genetics. By the late 1930s, the conservation movement in the Soviet Union had been completely decimated (Weiner 1988b, pp. 255–56).

The disconnection of Soviet thought from ecological issues, from the 1930s on, was severe and affected Marxism in the West as well, which

[10] Uranovsky was one of the first scientists to be arrested, in 1936, in the Stalinist purges (Medvedev [1971] 1989, p. 441). Accompanying Bukharin as a member of the Soviet delegation to the Second International Conference of the History of Science and Technology, London 1931, was also the brilliant plant geneticist N. I. Vavilov (one of the greatest figures in the history of ecological science), founder and first president of the Lenin Agricultural Academy, who applied a materialist method to the question of the origins of agriculture with the support of early Soviet science (Vavilov 1971). Like Bukharin and Uranovsky, he fell prey to the Stalinist purges.

between the 1930s and the 1970s tended to ignore ecological issues, though there was a revival of interest in this area in Marxism as well with the renewal of environmentalism following the publication of Rachel Carson's *Silent Spring* in 1962. To be sure, when Western Marxism had first emerged as a distinct tradition in the 1920s and 1930s, one of the major influences was the Frankfurt School, which developed an ecological critique (Horkheimer and Adorno 1972). But this critique was largely philosophical, and while it recognized the ecological insights in Marx's *Economic and Philosophical Manuscripts*, it lost sight of the ecological argument embedded in *Capital*. Hence, it generally concluded that classical Marxism (beginning with the later Marx) supported a "Promethean" philosophy of the straightforward domination of nature. Not until the 1960s and 1970s did a more complex interpretation begin to emerge in the writings of the thinkers influenced by the Frankfurt tradition (Schmidt 1971; Leiss 1974). And it was not until the late 1980s and 1990s that scholars began to resurrect Marx's argument on soil fertility and organic recycling (Perelman 1988; Hayward 1994; Foster 1997; Fischer-Kowalski 1997). Much of the renewed emphasis on Marx's (and Liebig's) treatment of soil fertility and its ecological implications has come from agronomists and ecologists concerned directly with the debates around the evolution of soil science and the struggles over agribusiness versus organic agriculture (Mayumi 1991; Magdoff, Lanyon, and Liebhardt 1997; Gardner 1997).

It is scarcely surprising, then, that interpretations of Marx within sociology, and environmental sociology in particular, have been affected by an "appropriation problem." Sociologists in general tend to have little knowledge of volume 3 of Marx's *Capital*, where his critique of capitalist agriculture (and of the undermining of soil fertility) is most fully developed, and while these issues were well-known to the generations of Marxist thinkers who immediately followed Marx, they largely vanished within Marxist thought in the 1930s. Even today, treatments of Marx's relation to ecology that purport to be comprehensive focus on his early writings, largely ignoring *Capital* (Dickens 1992). This appropriation problem had important ramifications. It left the appearance that there were no explicit linkages between human society and the natural world within classical Marxism, thus facilitating the notion that there was an unbridgeable gulf between classical sociology and environmental sociology.

Analogous appropriation problems might be raised with respect to the other classical theorists. Martinez-Alier (1987, pp. 183–92) has argued that Weber's important essay on Ostwald's social energetics has also been neglected; indeed it has yet to be translated into English. This has left the false impression that Weber had nothing to say in this area. Durkheim discussed the sociological origins of the classification of nature within what he called the "first philosophy of nature," and related this to modern

scientific evolutionism. He also commented in profound ways about Darwinian evolutionary theory, the indestructibility of matter, the conservation of energy, and so on (Durkheim and Mauss 1963, pp. 81–88; Durkheim [1893] 1984, pp. 208–9; Durkheim [1911–12] 1983, pp. 21–27, 69–70). The systematic character of his more naturalistic thinking has never been properly addressed, and works like *Pragmatism and Sociology*, in which he presents some of his more complex views in this regard, have generally been ignored. Nevertheless, it is clear that his analysis pointed toward a complex, coevolutionary perspective. "Sociology," he wrote, "introduces a relativism that rests on the relation between the physical environment on the one hand and man on the other. The physical environment presents a relative fixity. It undergoes evolution, of course; but reality never ceases to be what it was in order to give way to a reality of a new kind, or to one constituting new elements. .... The organic world does not abolish the physical world and the social world has not been formed in contradistinction to the organic world, but together with it" (Durkheim 1983, pp. 69–70).

## The Definitional Problem

Along with the appropriation problem, which deals with how received sociology has been affected by the selective appropriation of the classical tradition, there is also the definitional problem, which stands for the fact that sociology's (specifically environmental sociology's) failure to address the classical inheritance in this regard is at least partly due to overly narrow, preconceived definitions as to what constitutes genuinely environmental thought.

Here a major role was assumed by the contrast, drawn by Catton and Dunlap (1978), between the "human exemptionalist paradigm" and the "new environmental paradigm." All of the competing perspectives in sociology, such as "functionalism, symbolic interactionism, ethnomethodology, conflict theory, Marxism, and so forth" were seen as sharing a common trait of belonging to a "human exceptionalist paradigm" (later renamed "human exemptionalist paradigm"), and thus the "apparent diversity" of these theories was "not as important as the fundamental anthropocentrism underlying *all* of them" (Catton and Dunlap 1978, p. 42). The human exemptionalist paradigm was depicted as embracing the following assumptions: (1) the existence of culture makes human beings unique among the creatures of the earth, (2) culture evolves much more rapidly than biology, (3) most human characteristics are culturally based and hence can be socially altered, and (4) a process of cultural accumulation means that human progress can be cumulative and without limit. The habits of mind produced by this human exemptionalist paradigm,

Catton and Dunlap (1978, pp. 42–43) argued, led to an overly optimistic faith in human progress, a failure to acknowledge ecological scarcity, and a tendency to neglect fundamental physical laws such as the entropy law.

For Catton and Dunlap, this "human exemptionalist paradigm," which encompassed nearly all of existing sociology could be contrasted to what they termed the "new environmental paradigm" emerging from environmental sociology, which was based on the following assumptions: (1) human beings are one of many species that are interdependently connected within the biotic community; (2) the biotic community consists of an intricate web of nature, with complex linkages of cause and effect; and (3) the world itself is finite, there are natural (physical, biological) limits to social and economic progress (1978, p. 45). In contrast to the "anthropocentrism" that characterized the human exemptionalist paradigm, the new environmental paradigm represented a shift toward what is now called an "ecocentric" point of view in which human beings are seen as part of nature, interconnected with other species and subject to the natural limits of the biosphere.

Ironically, the chief problem with this contrast between the human exemptionalist paradigm and the new environmental paradigm is that, even while emphasizing environmental factors, it tended to perpetuate a dualistic view of society versus the physical environment, anthropocentrism versus ecocentrism, and thus easily fell into the fallacy of the excluded middle (or a false dichotomy). There is a tendency in this view to see any theory that emphasizes socioeconomic progress or cultural accumulation as thereby "anthropocentric" and opposed to an "ecocentric" perspective, which seeks to decenter the human world and human interests. Nevertheless, logic suggests that there is no reason for such a stark opposition, since there are numerous ways in which sociology can embrace a concern for ecological sustainability without abandoning its emphasis on the development of human culture and production. Moreover, extreme ecocentrism runs the risk of losing sight of the sociological construction of much of the "natural world." Although classical sociology may have been anthropocentric to some extent in its focus on socioeconomic advance and its relative neglect of external nature, it was not necessarily antiecological (in the sense of ignoring natural limits) insofar as it acknowledged ecological sustainability as a requirement of social progress. The current preoccupation with sustainable development and coevolutionary theories within environmental discussions suggests that there have always been complex views that attempted to transcend the dualisms of humanity versus nature, anthropocentrism versus ecocentrism, socioeconomic progress versus natural limits.

Marx in particular has been criticized for being "anthropocentric" rather than "ecocentric" in orientation and hence outside of the framework of green theory (Eckersley 1992, pp. 75–95). Yet this kind of dualistic

conception would have made little sense from his more dialectical perspective, which emphasized the quality (and sustainablilty) of the *interaction* between society and its natural conditions. It is the commitment to ecological sustainability, not the abstract notion of "ecocentrism," which most clearly defines whether a theory is part of ecological discourse. Moreover, a comprehensive *sociology* of the environment must by definition be co-evolutionary in perspective, taking into account changes in both society and nature and their mutual interaction.

## CONCLUSION: THE ELEMENTS OF ENVIRONMENTAL SOCIOLOGY

The burden of argument in this article has been to demonstrate, using the case of Marx, that it is wrong to contend that classical sociology "was constructed as if nature didn't matter" (Murphy 1996, p. 10). A central claim of this article, backed up by logic and evidence, has been that each of the six ecological blinders commonly attributed to Marx—namely his alleged inability to perceive (1) the exploitation of nature, (2) nature's role in the creation of wealth, (3) the existence of natural limits, (4) nature's changing character, (5) the role of technology in environmental degradation, and (6) the inability of mere economic abundance to solve environmental problems—are in fact wrongly (or misleadingly) attributed to him. The point of course is not that Marx provided definitive treatments of all of these problems but rather that he was sufficiently cognizant of these issues to elude the main traps and to work the vitally important notion of the "human metabolism with nature" into his overall theoretical framework. Hence his work constitutes a possible starting point for a comprehensive sociology of the environment. No doubt some will still insist, despite the argument presented above, that Marx did not place sufficient *emphasis* on natural conditions, or that his approach was too anthropocentric, more along the lines of utilitarian-conservationism that genuine green radicalism. Some will still say that he in fact never entirely renounced economic development despite his insistence on a sustainable relation to the earth. But the evidence regarding his concern with ecological issues—particularly the crisis of the soil as it was perceived in the mid-19th century—is too extensive, and too much a part of his overall critique of capitalism, to be simply disregarded. Marx certainly argued *as if nature mattered,* and his sociology thus takes on a whole new dimension when viewed from this standpoint.

Just as Marx translated his early theory of the alienation of labor into more material terms through his later analysis of exploitation and the degradation of work, so he translated his early notion of the alienation of nature (part of the Feuerbachian naturalism that pervaded his *Economic and Philosophical Manuscripts*) into more material terms through his later

concept of a metabolic rift. Without the latter concept, it is impossible to understand Marx's developed analysis of the antagonism of town and country, his critique of capitalist agriculture, or his calls for the "restoration" of the necessary metabolic relation between humanity and the earth, that is, his basic notion of sustainability. Marx's response to Liebig's critique of capitalist agriculture was coupled, moreover, with a sophisticated response to Darwin's evolutionary theory. What emerges from this is a historical materialism that is ultimately connected to natural history; one that rejects the crude, one-sided traditions of mechanical materialism, vitalism, and social Darwinism that existed in Marx's day. Yet, at the same time, Marx avoided falling into the trap of Engels's later "dialectical materialism," which, ironically, drew too heavily on both Hegel's *Logic* and his *Philosophy of Nature*, abstractly superimposing a despiritualized Hegelian dialectic (i.e., conceived in purely logical terms, divorced from Hegel's self-mediating spirit) on top of what was otherwise a mechanical view of the universe. Instead, Marx provides, as we have seen, a cautious constructionism, fully in tune with his own practical materialism, which always emphasized the role of human praxis, while remaining sensitive to natural conditions, evolutionary change, and the metabolic interaction of humanity and the earth.

Marx's main contribution in this area was methodological. He saw "the economic formation of society" as part of a process of "natural history" and struggled within his critique of political economy to take account of both natural conditions and the human transformation of nature (1976, p. 92). In the process, he applied a dialectical mode of analysis not to external nature itself (recognizing that the dialectic had no meaning aside from the self-mediating role of human beings as the agents of history) but rather to the *interaction* between nature and humanity, emphasizing the alienation of nature in existing forms of reproduction and the contradictory, nonsustainable character of the metabolic rift between nature and society that capitalism in particular had generated. Moreover, Marx conceived this metabolic rift not simply in abstract terms but in terms of the concrete crisis represented by the degradation of the soil and by the problem of human and animal "wastes" that engulfed the cities. Both were equal indications, in his analysis, of the metabolic rift between humanity and the soil, reflected in the antagonism of town and country.

The way in which Marx's analysis prefigured some of the most advanced ecological analysis of the late 20th century—particularly in relation to issues of the soil and the ecology of cities—is nothing less than startling. Much of the recent work on the ecology of the soil (Magdoff et al. 1997; Mayumi 1991; Gardner 1997) has focused on successive, historical breaks in nutrient cycling. The first such break, associated with the second agricultural revolution, is often conceived in essentially the same

terms in which it was originally discussed by Liebig and Marx and is seen as related to the physical removal of human beings from the land. This resulted in the failure to recycle human organic wastes back to the land, as well as the associated break in the metabolic cycle and the net loss to the soil arising from the transfer of organic products (food and fiber) over hundreds and thousands of miles. It was these developments that made the creation of a fertilizer industry necessary. A subsequent break occurred with the third agricultural revolution (the rise of agribusiness), which was associated in its early stages with the removal of large animals from farms, the creation of centralized feedlots, and the replacement of animal traction with farm machinery. No longer was it necessary to grow legumes, which had the beneficial effect of naturally fixing nitrogen in the soil, in order to feed ruminant animals. Hence, the dependence on fertilizer nitrogen increased, with all sorts of negative environmental consequences, including the contamination of ground water, the "death" of lakes, and so on. These developments, and other related processes, are now seen as related to the distorted pattern of development that has characterized capitalism (and other social systems such as the Soviet Union that replicated this pattern of development, sometimes in even more distorted fashion), taking the form of a more and more extreme metabolic rift between city and country—between what is now a mechanized humanity and a mechanized nature. Similarly, the ecological problem of the city is increasingly viewed in terms of its metabolic relationship to its external environment (focusing on the flows of organic nutrients, energy, etc.) and the ecological distortions that this entails (Wolman 1965; Giradet 1997; Fischer-Kowalski 1997; Opschoor 1997).

The fact that Marx was able to conceive a sociological approach that pointed to these developments when they were still in their very early stages represents one of the great triumphs of classical sociological analysis. It stands as a indication of how sociology could be extended into the ecological realm. It reinforces the view that ecological analysis, devoid of sociological insight, is incapable of dealing with the contemporary crisis of the earth—a crisis which has its source and its meaning ultimately in society itself.

It is not just Marxist sociology that is in a position to draw on Marx's insights in this respect, which are sociological as much as they are Marxist. Moreover, other paradigms within classical sociology have much more to contribute to the analysis of the natural environmental context of human social development than is commonly supposed. There is no doubt that Weber and Durkheim were both concerned in their own ways with the metabolic interaction between nature and society. Although systematic investigations into the work of Weber and Durkheim in this respect still

have to be undertaken, it is not to be doubted that embedded in their sociologies were important insights into ecological problems. When Weber wrote at the end of *The Protestant Ethic and the Spirit of Capitalism* of a civilization characterized by "mechanized petrification" that might continue along the same course—that of formal or instrumental rationality—"until the last ton of fossilized coal" was burnt, he was suggesting the possibility of a wider social and environmental critique of this civilization (Weber [1904–5] 1930, pp. 181–82). Likewise, Durkheim's discussions of Darwinian theory and its implications for social analysis pointed the way toward a sociological understanding of the coevolution of nature and society. In the cases of Weber and Durkheim—as in Marx—we may surmise that an appropriation problem, coupled with a definitional problem, has hindered the appreciation of the way in which their sociologies took natural conditions into account.

Today, even among leading environmental sociologists who criticized the classical traditions of sociology for failing to take into account the physical environment, there is a dawning recognition that these classical traditions have proven themselves to be resilient in the face of challenges of environmental sociologists and are open to reinterpretation and reformulation along lines that give greater weight to ecological factors. Dunlap points to the emergence, in recent years, of "'greener' versions of Marxist, Weberian and symbolic interactionist theories" (1997, p. 34). Ironically, it is coming to be recognized that the problem of "human exemptionalism," that is, the neglect of the physical environment, may have been less characteristic of classical sociology than it was of the sociology that predominated after World War II—during a period when the faith in technology and the human "conquest" of nature reached heights never before attained, only to lead to disillusionment and crisis beginning with the 1960s. Developing an environmental sociology as an integral part of sociology as a whole thus requires that we reach back into past theories in order to develop the intellectual means for a thoroughgoing analysis of the present. For environmental sociology the crucial issue today is to abandon the "strong constructionism" of most contemporary sociological theory, which tends to view the environment as simply a product of human beings, and to move toward a more "cautious constructionism" that recognizes that there is a complex metabolic relation between human beings and society (Dunlap 1997, pp. 31–32, 35; Dickens 1996, p. 71). Surprisingly, this is turning out to be an area in which the classical sociology of the mid-19th and early 20th centuries still has much to teach us as we enter the 21st century—a century that is bound to constitute a turning point for good or ill in the human relation to the environment.

REFERENCES

Altvater, Elmar. 1993. *The Future of the Market.* London: Verso.
Bacon, Francis. (1620) 1994. *Novum Organum.* Chicago: Open Court.
Bailes, Kendall. 1990. *Science and Russian Culture in an Age of Revolutions.* Bloomington: Indiana University Press.
Beck, Ulrich. 1995. *Ecological Enlightenment.* Atlantic Highlands, N.J.: Humanities Press.
Benton, Ted. 1989. "Marxism and Natural Limits." *New Left Review* 178:51–86.
———. 1994. "Biology and Social Theory in the Environmental Debate." Pp. 28–50 in *Social Theory and the Global Environment,* edited by Michael Redclift and Ted Benton. New York: Routledge.
Benton, Ted, and Michael Redclift. 1994. "Introduction." Pp. 1–27 in *Social Theory and the Global Environment,* edited by Michael Redclift and Ted Benton. New York: Routledge.
Bing, Franklin C. 1971. "The History of the Word 'Metabolism.'" *Journal of the History of Medicine and Allied Sciences* 26 (2): 158–80.
Brock, William H. 1997. *Justus von Liebig.* Cambridge: Cambridge University Press.
Bukharin, Nikolai. 1925. *Historical Materialism: A System of Sociology.* New York: International.
———. 1971. "Theory and Practice from the Standpoint of Dialectical Materialism." Pp. 11–33 in *Science at the Crossroads.* London: Frank Cass
Burch, William. 1971. *Daydreams and Nightmares.* New York: Harper & Row.
Burkett, Paul. 1997. "Nature in Marx Reconsidered." *Organization and Environment* 10 (2): 164–83.
Buttel, Frederick. 1986. "Sociology and the Environment." *International Social Science Journal* 109:337–56.
———. 1987. "New Directions in Environmental Sociology." *Annual Review of Sociology* 13:465–88.
———. 1996. "Environmental and Resource Sociology." *Rural Sociology* 61 (1): 56–76.
Caneva, Kenneth. 1993. *Robert Mayer and the Conservation of Energy.* Princeton, N.J.: Princeton University Press.
Carey, Henry. 1858. *Letters to the President on the Foreign and Domestic Policy of the Union.* Philadelphia: M. Polock.
———. (1858–59) 1867. *Principles of Social Science,* vol. 2. Philadelphia: J. B. Lippincott.
———. (1847) 1967a. *The Past, the Present and the Future.* New York: Augustus M. Kelley.
———. (1853) 1967b. *The Slave Trade Domestic and Foreign.* New York: Augustus M. Kelley.
Catton, William. 1982. *Overshoot.* Urbana: University of Illinois Press.
Catton, William, and Riley Dunlap. 1978. "Environmental Sociology: A New Paradigm." *American Sociologist* 13:41–49.
Churchill, Ward. 1996. *From a Native Son.* Boston: South End.
Clark, John. 1989. "Marx's Inorganic Body." *Environmental Ethics* 11:243–58.
Cohen, Stephen. 1980. *Bukharin and the Bolshevik Revolution.* Oxford: Oxford University Press.
Darwin, Charles. (1859) 1968. *The Origin of Species.* Middlesex: Penguin.
Deléage, Jean-Paul. 1994. "Eco-Marxist Critique of Political Economy." Pp. 37–52 in *Is Capitalism Sustainable?* edited by Martin O'Connor. New York: Guilford.
Dickens, Peter. 1992. *Society and Nature.* Philadelphia: Temple University.
———. 1996. *Reconstructing Nature.* New York: Routledge.
———. 1997. "Beyond Sociology." Pp. 179–92 in *International Handbook of Environ-*

*mental Sociology,* edited by Michael Redclift and Graham Woodgate. Northampton, Mass.: Edward Elgar.

Dryzek, John. 1997. *The Politics of the Earth.* Oxford: Oxford University Press.

Dunlap, Riley. 1997. "The Evolution of Environmental Sociology." Pp. 21–39 in *International Handbook of Environmental Sociology,* edited by Michael Redclift and Graham Woodgate. Northampton, Mass.: Edward Elgar.

Dunlap, Riley, and William Catton. 1979. "Environmental Sociology." Pp. 57–85 in *Progress in Resource Management and Environmental Planning,* vol. 1. Edited by Timothy O'Riordan and Ralph D'Arge New York: John Wiley & Sons.

———. 1994. "Struggling with Human Exceptionalism." *American Sociologist* 25 (1): 5–30.

Dunlap, Riley, and Kenneth Martin. 1983. "Bringing Environment into the Study of Agriculture." *Rural Sociology* 48 (2): 201–18.

Durkheim, Émile. (1911–12) 1983. *Pragmatism and Sociology.* Cambridge: Cambridge University Press.

———. (1893) 1984. *The Division of Labor in Society.* New York: Free Press.

Durkheim, Émile, and Marcel Mauss. 1963. *Primitive Classification.* Chicago: University of Chicago Press.

Eckersley, Robyn. 1992. *Environmentalism and Political Theory.* New York: State University of New York Press.

Engels, Friedrich. (1874–80) 1940. *The Dialectics of Nature.* New York: International.

——— 1964. "Outlines of a Critique of Political Economy." Pp. 197–226 *The Economic and Philosophic Manuscripts of 1844,* edited by Dirk J. Struik. New York: International.

———. (1872) 1975. *The Housing Question.* Moscow: Progress.

Ernle, Lord. (1912) 1961. *English Farming Past and Present.* Chicago: Quadrangle.

Ferkiss, Victor. 1993. *Nature, Technology and Society.* New York: New York University Press.

Feshbach, Murray, and Arthur Friendly, Jr. 1992. *Ecocide in the U.S.S.R.* New York. Basic.

Fischer-Kowalski, Marina. 1997. "Society's Metabolism." Pp. 119–37 in *International Handbook of Environmental Sociology,* edited by Michael Redclift and Graham Woodgate. Northampton, Mass.· Edward Elgar.

Foster, John Bellamy. 1997. "The Crisis of the Earth." *Organization and Environment* 10 (3): 278–95

Gardner, Gary. 1997. *Recycling Organic Wastes.* Washington, D.C.: Worldwatch.

Geertz, Clifford. 1963. *Agricultural Involution.* Berkeley: University of California Press.

Georgescu-Roegen, Nicholas. 1971. *The Entropy Law in the Economic Process.* Cambridge, Mass.: Harvard University Press.

Giddens, Anthony. 1981. *A Contemporary Critique of Historical Materialism.* Berkeley and Los Angeles: University of California Press.

Giradet, Herbert. 1997. "Sustainable Cities." *Architectural Design* 67:9–13.

Goldblatt, David. 1996. *Social Theory and the Environment.* Boulder, Colo.: Westview.

Golley, Frank 1993. *A History of the Ecosystem Concept in Ecology.* New Haven, Conn.: Yale University Press.

Gould, Stephen Jay. 1987 *An Urchin in the Storm.* New York: W. W. Norton.

Grundmann, Reiner. 1991. *Marxism and Ecology.* Oxford: Oxford University Press.

Harvey, David. 1996. *Justice, Nature and the Geography of Difference.* New York: Blackwell.

Hayward, Tim. 1994. *Ecological Thought.* Cambridge, Mass.: Polity.

Hillel, Daniel. 1991. *Out of the Earth.* Berkeley and Los Angeles: University of California.

Horkheimer, Max, and Theodor Adorno. 1972. *The Dialectic of Enlightenment.* New York: Continuum.

Jacobs, Michael. 1994. "The Limits to Neoclassicism." Pp. 67–91 in *Social Theory and the Environment,* edited by Michael Redclift and Ted Benton. New York: Routledge.

Järvikoski, Timo. 1996. "The Relation of Nature and Society in Marx and Durkheim." *Acta Sociologica* 39 (1): 73–86.

Johnston, James. 1851. *Notes on North America.* London: William Blackwood & Sons.

Kautsky, Karl. (1899) 1988. *The Agrarian Question,* 2 vols. Winchester, Mass.: Zwan.

Komarov, V. L. 1935. "Marx and Engels on Biology." Pp. 190–234 in *Marxism and Modern Thought.* New York: Harcourt, Brace.

Lebowitz, Michael. 1992. *Beyond Capital.* London: Macmillan.

Leiss, William. 1974. *The Domination of Nature.* Boston: Beacon.

Lenin, V. I. 1961. *Collected Works,* vol. 5. Edited by Victor Jerome. Moscow: Progress.

Liebig, Justus von. 1859. *Letters on Modern Agriculture.* London: Walton & Maberly

———. 1863. *The Natural Laws of Husbandry.* New York: D. Appleton

———. 1865. *Letters on the Subject of the Utilization of the Metropolitan Sewage.* London· W. H. Collingridge.

———. (1842) 1964. *Animal Chemistry or Organic Chemistry in its Application to Physiology and Pathology.* New York: Johnson Reprint.

Magdoff, Fred, Less Lanyon, and Bill Liebhardt. 1997. "Nutrient Cycling, Transformations and Flows." *Advances in Agronomy* 60:1–73.

Malthus, Thomas. 1970. *Pamphlets.* New York: Augustus M. Kelley.

Margulis, Lynn, et. al. 1998. Foreword to *The Biosphere,* by V. I. Vernadsky. New York: Copernicus.

Marshall, Alfred. 1920. *Principles of Economics.* London: Macmillan.

Martinez-Alier, Juan. 1987. *Ecological Economics.* Oxford: Basil Blackwell.

Marx, Karl. (1847) 1963. *The Poverty of Philosophy.* New York: International.

———. (1875) 1971a. *Critique of the Gotha Programme.* Moscow: Progress.

———. (1861–63) 1971b. *Theories of Surplus Value,* pt. 3. Moscow: Progress.

———. (1857–58) 1973. *Grundrisse.* New York: Vintage.

———. 1974. *Early Writings.* New York: Vintage.

———. 1975. *Texts on Method.* Oxford: Basil Blackwell.

———. (1867) 1976. *Capital,* vol. 1. New York: Vintage.

———. (1865–70) 1978. *Capital,* vol. 2. New York: Vintage.

———. (1863–65) 1981. *Capital,* vol. 3. New York Vintage.

Marx, Karl, and Friedrich Engels. (1848) 1967. *The Communist Manifesto.* New York: Monthly Review

———. 1975a. *Collected Works.* New York: International.

———. 1975b. *Selected Correspondence,* edited by S W. Ryazanskaya. Moscow: Progress.

Mayumi, Kozo. 1991. "Temporary Emancipation from the Land." *Ecological Economics* 4 (1): 35–56.

McConnell, Campbell. 1987. *Economics.* New York: McGraw Hill.

McLaughlin, Andrew. 1990. "Ecology, Capitalism, and Socialism." *Socialism and Democracy* 10:69–102.

———. 1993. *Regarding Nature.* Albany: State University of New York Press.

Medvedev, Roy. (1971) 1989. *Let History Judge.* New York: Columbia University Press.

Milton, Kay. 1996. *Environmentalism and Cultural Theory.* New York· Routledge.

Murphy, Raymond. 1994. *Rationality and Nature.* Boulder, Colo.: Westview.

———. 1996. *Sociology and Nature.* Boulder, Colo.: Westview.

Norgaard, Richard. 1994. *Development Betrayed.* New York: Routledge.

Nove, Alec. 1987. "Socialism." Pp. 398–407 in *The New Palgrave Dictionary of Eco-*

*nomics,* vol. 4. Edited by John Eatwell, Murray Milgate, and Peter Newman. New York: Stockton.

O'Connor, James. 1998. *Natural Causes.* New York. Guilford.

Odum, Eugene. 1969. "The Strategy of Ecosystem Development" *Science* 164:262–70.

Opschoor, J. B. 1997. "Industrial Metabolism, Economic Growth and Institutional Change." Pp. 274–86 in *International Handbook of Environmental Sociology,* edited by Michael Redclift and Graham Woodgate. Northampton, Mass.: Edward Elgar.

Pannekoek, Anton. 1912. *Marxism and Darwinism.* Chicago: Charles H. Kerr.

Parsons, Howard, ed. 1977. *Marx and Engels on Ecology.* Westport, Conn.: Greenwood.

Perelman, Michael. 1988. "Marx and Resources." *Environment, Technology and Society* 51:15–19.

———. 1993. "Marx and Resource Scarcity." *Capitalism, Nature, Socialism* 4 (2): 65–84.

Peterson, D. J. 1993. *Troubled Lands.* Boulder, Colo.: Westview.

Pigou, A. C. 1920 *The Economics of Welfare.* London: Macmillan.

Quaini, Massimo. 1982. *Marxism and Geography.* Totowa, N.J.: Barnes & Noble.

Redclift, Michael. 1984. *Development and the Environmental Crisis.* New York: Methuen.

Redclift, Michael, and Graham Woodgate. 1994. "Sociology and the Environment." Pp. 51–66 in *Social Theory and the Global Environment,* vol. 1. Edited by Michael Redclift and Ted Benton. New York: Routledge.

Ricardo, David. 1951. *Principles of Political Economy and Taxation.* Cambridge: Cambridge University Press.

Sauer, Carl. 1963. *Land and Life.* Berkeley. University of California Press.

Schmidt, Alfred. 1971. *The Concept of Nature in Marx.* London: New Left.

Schnaiberg, Allen. 1980. *The Environment.* Oxford: Oxford University Press.

Shanin, Teodor. 1983. *Late Marx and the Russian Road.* New York: Monthly Review.

Skaggs, J. M. 1994. *The Great Guano Rush.* New York: St. Martin's.

Thompson, F. M. L. 1968. "The Second Agricultural Revolution, 1815–1880." *Economic History Review* 21 (1): 62–77.

Uranovsky, Y. M. 1935. "Marxism and Natural Science." Pp. 136–74 in *Marxism and Modern Thought.* New York: Harcourt, Brace.

Vavilov, N. I. 1971. "The Problem of the Origin of the World's Agriculture in the Light of the Latest Investigations." Pp. 97–106 in *Science at the Crossroads.* London: Frank Cass.

Vernadsky, V. I. (1922) 1998. *The Biosphere.* New York: Copernicus.

Weber, Max. (1904–5) 1930. *The Protestant Ethic and the Spirit of Capitalism.* London: Unwin Hyman.

Weiner, Douglas. 1988a. *Models of Nature.* Bloomington: Indiana University Press.

———. 1988b. "The Changing Face of Soviet Conservation." Pp. 252–73 in *The Ends of The Earth,* edited by Donald Worster. New York: Cambridge University Press.

Wolman, Abel. 1965. "The Metabolism of Cities." *Scientific American* 213 (3): 179–90.

World Commission on Environment and Development. 1987. *Our Common Future.* New York: Oxford University Press

# The Endogeneity of Legal Regulation: Grievance Procedures as Rational Myth[1]

Lauren B. Edelman
*University of California—Berkeley*

Christopher Uggen
*University of Minnesota*

Howard S. Erlanger
*University of Wisconsin—Madison*

Most accounts of organizations and law treat law as largely exogenous and emphasize organizations' responses to law. This study proposes a model of endogeneity among organizations, the professions, and legal institutions. It suggests that organizations and the professions strive to construct rational responses to law, enabled by "rational myths" or stories about appropriate solutions that are themselves modeled after the public legal order. Courts, in turn, recognize and legitimate organizational structures that mimic the legal form, thus conferring legal and market benefits upon organizational structures that began as gestures of compliance. Thus, market rationality can follow from rationalized myths: the professions promote a particular compliance strategy, organizations adopt this strategy to reduce costs and symbolize compliance, and courts adjust judicial constructions of fairness to include these emerging organizational practices. To illustrate this model, a case study of equal employment opportunity (EEO) grievance procedures is presented in this article.

The meaning of law regulating organizations unfolds dynamically across organizational, professional, and legal fields. Legislative action and judicial interpretation offer law its official stamp, but organizations—and the

larger environmental fields in which they operate—interact with legal institutions to create ritualized ideas about what constitutes a rational response to law. It is these *ideologies of rationality*—the accounts, stories, and myths about how organizations should respond to law—rather than the substantive law itself that ultimately determine both organizations' strategic responses to law and the courts' responses to organizational actions.

Law, then, should be understood more as a rhetorical and symbolic resource than as an articulate mandate (Stryker 1990, 1994). The construction of law is invariably contested: many voices contribute to the process of legal enactment and vie for favorable interpretations of law once it is enacted. The more ambiguous and politically contested the law, the more open it is to social construction. Law regulating organizations is especially open to social construction because the corporate lobby is usually successful in softening regulation that infringes on corporate interests, thus producing broad, vague mandates. Under such conditions, organizations actively participate in constructing the meaning of compliance, and this construction process generates ideologies of rationality, which legitimate and reinforce particular compliance strategies. That organizations are both responding to and constructing the law that regulates them renders law "endogenous"; the content and meaning of law is determined within the social field that it is designed to regulate.

In this article, we explore the evolving ideologies of rationality within the context of equal employment opportunity (EEO) law, most notably Title VII of the 1964 Civil Rights Act, which forbids discriminatory treatment on the basis of race, color, sex, religion, or national origin. We consider in particular the interactions among organizations, the legal and personnel professions, and the courts as they construct the meaning of compliance with EEO law. Since EEO law is a particularly ambiguous and controversial area, the effects of ideologies of rationality on forms of compliance are especially visible in this case. Nevertheless, since most forms of law regulating organizations use broad language, our arguments should be applicable to other legal areas, albeit to varying degrees.

In particular, we focus on the increasing portrayal of internal EEO grievance procedures as a "rational" mode of compliance with EEO law. EEO law simply mandates nondiscrimination; it is silent as to the question of what actions an employer might take to rebut an employee's claim of discrimination. Yet, grievance procedures have emerged over the past few decades as the primary symbol of nondiscrimination and as the most rational way for employers to insulate themselves from legal liability. We use

of Law and Society, University of California, 2240 Piedmont Avenue, Berkeley, California 94720  E-mail: ledelman@uclink4.berkeley.edu

three data sources to document the construction of ideologies of grievance procedure rationality: (1) the professional personnel literature to examine the claims of management, legal, and personnel professionals about grievance procedures; (2) survey data to study organizational creation of the procedures and their impact on complaints; and (3) federal court cases to illustrate the judiciary's construction of these procedures.

Our data suggest that the professions proffer stories about the legal value of grievance procedures to organizations and that the stories influence both organizational adoption of the procedures and legal considerations by courts. These stories are based on theories that have only tenuous legal foundation or greatly exaggerate the scope of court rulings. Nonetheless, as they are told and retold in the professional journals, the stories tend to become widely accepted in organizational fields and to influence ideas about organizational rationality across organizational, professional, and legal realms. Over time, actors in each of these realms come to equate grievance procedures with compliance, leading to the prevalent belief that organizations can avoid significant legal costs by creating these procedures.

So far, then, ours is a typically institutionalist argument. We also contend, however, that organizational ideologies of rationality *induce the judiciary* to incorporate grievance procedures into legal constructions of compliance with EEO law. Further, because of judicial decisions that reinforce ideologies of rationality, these procedures begin to confer tangible economic benefits on organizations. Thus what may have begun as somewhat of a stab in the dark—an educated guess modeled after the public legal order—acquires market rationality because of processes of institutionalization.

This article goes beyond previous analyses of organizational response to law (Edelman 1990, 1992; Edelman, Abraham, and Erlanger 1992; Dobbin et al. 1988; Dobbin et al. 1993; Sutton et al. 1994) in several ways. First, whereas previous studies have focused on the institutional sources of diffusion of legal structures, this study seeks to specify how the diffusion of institutional structures affects their market rationality, that is, the cost savings to organizations. Second, whereas previous studies refer to the *effects* of court decisions, we introduce courts (and to a lesser extent administrative agencies) as important players in the *production* of ideologies of rationality. Third, in contrast to institutional perspectives that do not empirically examine their assertion that institutionalized beliefs are mythical, we provide tests of the major claims about the rationality of grievance procedures that have become widely prevalent in organizational and professional communities. Finally, we trace the development of a particular ideology of rationality over time, which allows a stronger causal argument

with respect to the interdependent roles of organizations, the professions, and the judiciary.

In the next section, we review theoretical perspectives on organizational response to law and propose a synthesis of conflicting views of rationality in the institutional and more traditional rational perspectives. Then, our argument proceeds as follows: In the section on the socially constructed rationality of grievance procedures, we sample the professional personnel literature to illustrate what we believe to be a major source of the construction of rationality in organizational fields: arguments that EEO grievance procedures are a rational response to the threat of legal intervention and that they will insulate the organization from liability for discrimination. Using data from Edelman's national survey of organizational response to civil rights law, we show how these articles correspond to a dramatic increase in organizations' adoptions of internal grievance procedures.

Next, we turn to empirical analyses of the claims spread and legitimated by the personnel profession and show that those "rational myths" have little basis in fact, at least when they are first introduced. In the section on the myth of grievance procedures as insulators, we use data from the national survey of organizations to refute the empirical claim that the use of grievance procedures would result in fewer external complaints. Then, in the section on the myth of judicial deferral to grievance procedures, we show that the law in effect at the time the early articles were written offered little support for the claim that grievance procedures would insulate organizations from liability. However, we also show that over time, as defendants in employment discrimination cases increasingly used the grievance procedure defense, courts have become more likely to defer to organizations' grievance procedures and to consider them relevant to determinations of liability. Thus, in our conclusion, we argue that the process of institutionalization has started to come full circle: Even though grievance procedures were not important in legal doctrine when the personnel profession promoted them, their significance as evidence of fair treatment has increasingly become an accepted part of the legal arsenal in discrimination lawsuits, and courts are far more likely today than they were in the past to find organizations with EEO grievance procedures not liable for discrimination. Our story, then, provides an example of the reciprocal relationship between organizational behavior and legal regulation.

## THEORETICAL PERSPECTIVES

Legal rules are not self-enforcing. When a new legal rule is announced, those subject to it must determine what constitutes compliance and what

actions they will take to demonstrate compliance (Edelman 1992). In addressing how organizations respond to law, two important analytic frameworks are a market approach, which emphasizes rational adaptation to a set of external market conditions, and an institutional approach, which emphasizes normative cues that emanate from organizations' environments. We suggest that these apparently oppositional explanations are to some extent complementary and reinforcing when the endogeneity of law is observed over time.

Market approaches emphasize organizational agency and generally hold that economic markets reward efficiency and rationality in organizational structure. These theories tend to assume that efficiency is objectively knowable and largely a product of market conditions. Because rational perspectives focus on organizations' market behaviors, law has played a relatively minor role in these analyses. And classical economic approaches address EEO law but tend to view it as an impediment to market efficiency (Epstein 1992; Posner 1986; Becker 1971). Two more recent approaches emphasize organizational adaptation to the external environment: Resource dependence theory suggests that within the constraints set by law, organizations will adopt structures that minimize costs and maximize resource flows (Pfeffer and Salancik 1978), and transaction cost approaches view law as important in shaping the relative costs of governance through contractual or hierarchical arrangements (Williamson 1979, 1981). In these accounts, then, law is one of many market forces that may affect what course of action will offer the greatest financial success. Law is understood as an exogenous phenomenon, and organizations are assumed to be rational actors with good information, attempting to operate efficiently with respect to legal constraints.

Institutional theories challenge the notion of an objective rationality, arguing that concepts of rationality are socially constructed by nonmarket factors such as widely accepted norms and patterns of behavior (Meyer and Rowan 1977; DiMaggio and Powell 1983; Scott 1983; Powell and DiMaggio 1991). Institutional accounts derive from the phenomenological work of Berger and Luckmann (1967) and highlight the taken-for-granted nature of rules. Further, these accounts hold that organizational behavior is largely given by "rational myths" or belief systems that embody stories about cause and effect and successful solutions to problems (Meyer and Rowan 1977; Zucker 1977, 1988). These belief systems appear *rational* in that they specify in a rulelike manner what organizations must do to be efficient, but they are *myths* in that their efficacy depends on the fact that they are widely shared rather than inherently correct (Scott 1987). In this view, organizations lack agency because they are shaped to such a great extent by institutionalized rules. Other institutional accounts (e.g., DiMag-

gio 1988; Powell 1985), however, and in particular institutional studies that directly address the relation between law and organizations (e.g., Edelman 1990, 1992; Konrad and Linnehan 1995; Stryker 1994; see Suchman and Edelman [1996] for a review), tend to reject the phenomenological assumptions of the early institutionalism. These works continue to emphasize the responsiveness of organizations to their institutional environments but allow room for agency in the context of organizations' attempts to respond strategically to legal mandates. Sewell (1992) argues more generally that institutionalized schema provide resources for human agency.

The extant literature on law and organizations, then, raises the possibility of agency in organizational response to law but does not specify how institutional processes and agency might coexist. This article elaborates the institutional work on law and organizations by suggesting that organizational agency and institutional processes coexist through a process of *local rationality within the context of global institutionalism:* organizations seek to act rationally in response to law—to minimize costs and maximize resource flows—but the definition of rationality is constructed and evolves at the environmental level, driven by institutionalized stories about the value of particular organizational structures and actions.

There are two tenets to our argument: First, *organizations seek to act rationally* but are constrained by their institutional environment. For example, organizations adopt EEO grievance procedures in an effort to reduce the potential for liability, but they choose this adaptation in preference to others because grievance procedures have acquired an aura of fairness and efficacy. Second, *institutional environments influence market rationality,* that is, the social valuation of institutionalized structures affects their benefits to organizations. In our example, the prevalence of grievance procedures and the arguments of professionals render courts more likely to consider grievance procedures as reasonable steps toward fair treatment. Thus, the adoption of institutionalized structures often provides monetary benefits, as well as legitimacy, to organizations. This approach is decidedly institutional in that we see rationality as ultimately socially constructed. But it explicitly incorporates elements of resource dependence and transaction costs analysis by recognizing organizational efforts to adapt strategically to their legal environments (compare Oliver 1991) and by suggesting that social constructions, when recognized by the courts, render certain organizational behaviors objectively rational. Thus, institutional processes work in concert with, rather than in place of, market rationality. Over time, rational myths can become, or can at least influence, market reality.

THE SOCIALLY CONSTRUCTED RATIONALITY OF GRIEVANCE
PROCEDURES

In this section, we address the construction of the rationality of EEO
grievance procedures during the first two decades following the passage
of the 1964 Civil Rights Act. During this period, organizations were con-
fronted with ambiguous legal mandates and the problem of defining com-
pliance (see Edelman 1992). EEO grievance procedures gained popularity
as organizations looked to one another and to the personnel and legal
professions for rational solutions to this problem.

### The Professions' Proposal of Grievance Procedures as the Solution

Professional networks generally play an important role in the diffusion of
new forms of governance and responses to law (DiMaggio and Powell
1983; Baron, Dobbin, and Jennings 1986; Edelman et al. 1992; Jacoby
1985; Suchman 1993) and, not surprisingly, appear to have been critical
in bringing grievance procedures into the legal consciousness of the orga-
nizational community. These networks consist of professional associations
and their conventions, commercial workshops on how to comply with law,
professional communications on the World Wide Web, and a large num-
ber of professional personnel journals.

To examine the professional construction of the rationality of EEO
grievance procedures, we reviewed 85 articles in the business literature
from 1964 to 1989 (personnel and management journals and law journals
relevant to businesses), addressing nonunion grievance procedures in the
context of EEO law.[2] The authors of these articles are mostly personnel
managers, management academics, or staff writers, although some are
lawyers. All are intended, however, for a personnel audience.[3]

The articles frame the rationality of grievance procedures in several
ways, which vary somewhat over time. In the late 1960s, articles that
mentioned grievance procedures tended to focus on their benefits for
avoiding unions (e.g., Corzine 1967), but by the early 1980s, the profes-
sions had turned their attention to avoiding legal problems and in particu-

[2] The articles were collected by the first author as part of her larger project on organi-
zations' EEO practices. An undergraduate assistant searched all issues of 10 major
professional personnel journals during 1964–89 and identified articles addressing
grievance procedures. For this study, we chose the articles that addressed grievance
procedures in the context of EEO law.

[3] Although some are written about the value of general grievance procedures applied
to EEO matters rather than about EEO grievance procedures specifically, the quota-
tions below show the types of claims that members of the personnel, legal, and man-
agement professions make about the rationality of grievance procedures generally.

lar to compliance with civil rights law. During this period, the articles emphasize (and arguably inflate) the threat of civil rights lawsuits and claim that organizations can substantially reduce both the number of such lawsuits and the potential for liability should lawsuits occur by creating internal grievance procedures. The articles generally hold that grievance procedures reduce lawsuits and complaints to external agencies by resolving claims internally, reduce liability by demonstrating compliance, reduce discrimination by providing information and locating management problems, reduce the appeal of unions by providing due process without dues, and raise morale by giving employees voice.[4]

Two specific assertions about the legal value of grievance procedures emerged during this period: first, internal grievance procedures help to resolve employee complaints so that employees would have no need to file charges with external EEO agencies; and second, if external complaints were filed, courts would look favorably on organizations that had taken steps to provide internal due process. The latter claim was made most often in connection with discussions of sexual harassment claims. For example, a 1981 article in *Personnel,* written by a lawyer and a management professor, begins with a blurb above the title stating, "Employers can protect themselves against liability for sexual harassment charges with a strong policy against such activity and a grievance procedure that expedites the processing of such complaints" (Linenberger and Keaveny 1981, p. 60). The article goes on to state, "If the employer has no knowledge of the harassment, liability can be avoided if two conditions have been met: (1) The employer has a policy discouraging sexual harassment, and the employee failed to use an existing grievance procedure; and (2) the sexually harassing situations are rectified as soon as the employer becomes aware of them" (Linenberger and Keaveny 1981, p. 60). In fact, as we point out below, there was very little legal support for this argument in 1981.

The personnel journals emphasize the advantages of internal grievance procedures for insulating organizations from liability, but they also suggest internal benefits: Grievance procedures provide a sense of justice to employees and will therefore improve morale and productivity. David Ewing, a managing editor of the *Harvard Business Review,* ties these themes together in a 1982 article in that journal.

---

[4] It is not just articles written about grievance procedures that emphasize their value. Virtually all articles that address civil rights law or civil rights issues mention the value of grievance procedures; moreover, we have never encountered an article that recommends that employers not have a grievance procedure. Other forums, such as commercial workshops on human resources issues and Web sites addressing human resource audiences also frequently recommend the adoption of grievance procedures.

> An effective form of due process has various advantages for a company. It helps to clear the air so that rumors of an abusive discharge or unfair handling of an objector do not circulate and build up and sometimes, in the end, create a worse situation than the original event. It is valuable for morale. . . . Also, attorneys with whom I have talked believe that an in-company hearing procedure can be helpful in case the employee objector takes the company to court; that is, *evidence that an objector has been turned down in a fair hearing will be admitted into a legal proceeding.* (Ewing 1982, p. 121; emphasis added)

Articles frequently emphasize the significant cost savings available to organizations that institute internal grievance procedures. A 1984 article, written by a lawyer and published in *Management Review* contains the following advice: "A good grievance procedure keeps problems within the company. It encourages fair treatment. Most importantly, it deters employees from seeking representation by outsiders—unions, government agencies (like the EEOC or OSHA) or even lawyers. A problem settled in-house can save tens of thousands in litigation costs" (Panken 1984, p. 42).

A 1985 article also points to the value of grievance procedures for avoiding employee lawsuits in the first place. The author, an investigative reporter, writes, "If an employee can get a fair hearing through a grievance system, there is less chance this employee will be tempted to go to the courts" (Condon 1985, p. 72).

By the mid-1980s, then, the theme that grievance procedures could internalize disputes and gain favor with courts appears to be well established in the professional literature. Grievance procedures had become widely accepted as a rational form of compliance with civil rights law and as a rational form of managing internal disputes. The question of *why* the professions advocate particular solutions to legal problems is an important one that we do not address directly in this article. But the professions' literature offers a number of explanations, including professions' battles of jurisdictions (Abbott 1988; Larson 1977) and their efforts to establish power within organizations (Pfeffer 1981; Edelman et al. 1992).

### Organizations' Adoption of EEO Grievance Procedures

The rationalization of grievance procedures is not simply an artifact of the personnel and legal professions; it is evidenced by the practices of organizations as well. In this section, we use data from Edelman's nationwide survey of EEO practices, conducted in 1989, to examine the adoption patterns of EEO grievance procedures (see Edelman [1992] for a discussion of the sample and response rates). As shown in table 1, of 346 organizations, 107 (31%) had created grievance procedures explicitly designed

TABLE 1

DESCRIPTIVE STATISTICS ON THE CREATION OF DISCRIMINATION GRIEVANCE
PROCEDURES

|  | Government | College | Business | Total |
|---|---|---|---|---|
| No. (%) of organizations creating specific grievance procedures:* | | | | |
| Procedure ................... . .......... ................... | 22 (45) | 22 (44) | 63 (26) | 107 (31) |
| No procedure .................. .... .................. | 27 (55) | 28 (56) | 184 (74) | 239 (69) |
| Total ............ ..... ................. ... ..... .. .... | 49 | 50 | 247 | 346 |
| The creation of specific grievance procedures by time period.† | | | | |
| Before 1970 .. .. ...................... ... .. .. ........ | 2 (10) | 2 (10) | 10 (22) | 14 (16) |
| 1970–79 .......... .. . ..... .. ................. . ... .. .. . | 11 (55) | 11 (52) | 11 (24) | 33 (38) |
| 1980–90† ...................... ... ............ ........... . | 7 (35) | 8 (38) | 24 (53) | 39 (45) |
| Total .. .. .......................... ... ................... | 20 | 21 | 45 | 86 |
| No. (%) of organizations creating specific grievance procedures—mail sample only:‡ | | | | |
| Procedure ........................ ...... ................. | 16 (47) | 15 (48) | 29 (20) | 60 (29) |
| No procedure . .. ............... .. .. .. ... . ........ | 18 (53) | 16 (52) | 113 (80) | 147 (71) |
| Total ............ .. .. ...... ............... ...... .. .. ... | 34 | 31 | 142 | 207 |
| The creation of specific grievance procedures by time period—mail sample only:§ | | | | |
| Before 1970 . ....... ...... ................ . ... ...... | 1 (6) | 2 (13) | 3 (13) | 6 (11) |
| 1970–79 ...................... .. .. ..... ................... | 9 (56) | 7 (47) | 7 (29) | 23 (42) |
| 1980–89† .. ..... ......................................... | 6 (38) | 6 (40) | 14 (58) | 26 (47) |
| Total .............. ..... .. ... ............ ... ............. | 16 | 15 | 24 | 55 |

* Between 1987 and 1990, 16 organizations created procedures These organizations were coded as *not* having a procedure in analyses predicting complaints in 1986

† Dates of the creation of specific procedures were not available to 21 cases For the event history analyses, these cases are omitted

‡ Between 1987 and 1990, 12 organizations created procedures These organizations were coded as *not* having a procedure in analyses predicting complaints in 1986.

§ Dates of the creation of specific procedures were not available for five cases. These cases are omitted for the event history analyses

to handle discrimination-related complaints by 1989. Of the 86 for which creation dates were available, 14 (16%) were created before 1970, 33 (38%) were created between 1970 and 1979, and 39 (45%) were created between 1980 and 1989.[5]

[5] Fig. 1 also shows the creation patterns for the 207 organizations that responded to the follow-up mail survey. The mail follow-up is discussed in more detail later in the section on the myth of grievance procedures as insulators. Analyses of complaints shown in later tables are based on the mail sample only.

Figure 1 shows an integrated hazard plot, which illustrates the pattern of adoption of these procedures over time.[6] The plot shows little change in the formation rate until about 1976, a time of relatively strong civil rights enforcement, and reflects a rapid diffusion of these structures during the 1970s. The adoption rate remains fairly high during the 1980s despite the weaker civil rights enforcement of the Reagan era. This suggests that rational myths, as much as enforcement threats, were motivating the diffusion process by this period. In fact, there appears to be a second burst of EEO grievance procedure creation in the second half of the 1980s.

### Why Is the Grievance Procedure Solution Attractive to Organizations?

Market-based approaches to organizational behavior would attribute the evolution of EEO grievance procedures to the market efficiency of these practices; like the personnel professionals, they would tend to see these procedures as valuable for employee morale and for minimizing legal risk (Thompson 1967; Pfeffer and Salancik 1978; Williamson 1975). The institutional literature, however, would attribute the evolution of EEO grievance procedures to the construction of rational myths or stories about the rationality about these procedures (Meyer and Rowan 1977; Scott and Meyer 1983; DiMaggio and Powell 1983). One of the key features of rational myths is that they appear so obvious that no one questions their veracity; they just *seem* right.

Our contention is that these myths originate from models that have already been institutionalized in other social arenas but that, over time, they influence law and, hence, market forces. In the case of grievance procedures, the institutionalized source of the myth is the legal order. Grievance procedures appear rational because they *look like* the system of appeals available in the public legal process, a basic and well-institutionalized feature of a legitimate normative order (see Edelman 1990). Since civil rights law is essentially a mandate of fair treatment for employees, regardless of race, sex, religion, or national origin, it seems only *natural*—and it is taken for granted—that grievance procedures, as the symbolic embodiment of due process, would in fact provide fair treatment. And therefore, it is taken for granted that the courts, society, and employ-

---

[6] Integrated hazard functions provide a nonparametric estimate of the integral of the hazard rate. Although the adoption rate is not directly observable, the hazard rate can be calculated based on the proportion of organizations at risk of experiencing an event at any given point in time that do in fact experience the event. The area under the curve represents the cumulative probability that organizations will create EEO grievance procedures. Changes in the slope of the center line reflect changes in the rate of EEO grievance procedure creation. The upper and lower lines represent the 95% confidence interval.
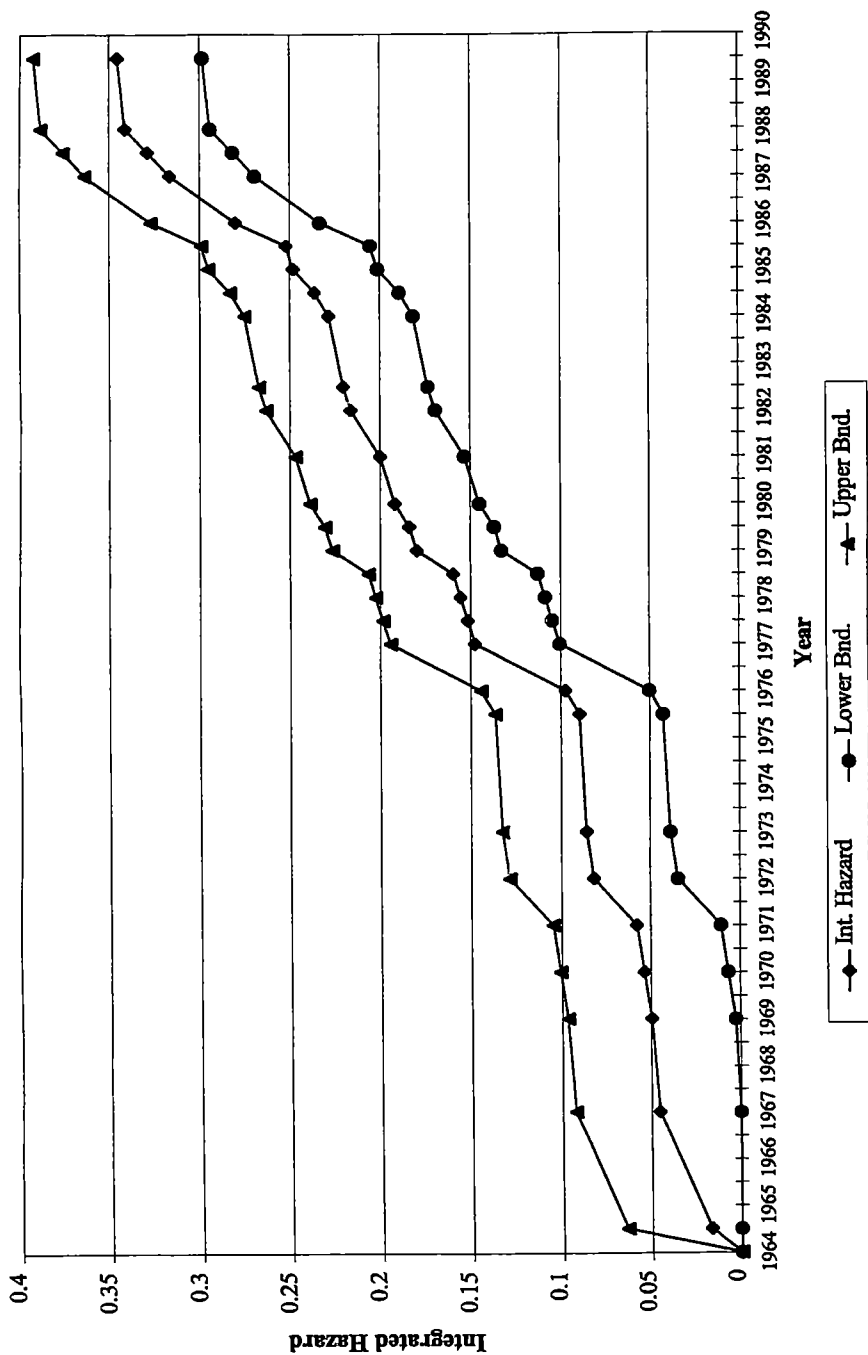
Fig. 1.—Creation of specific EEO grievance procedures

ees would in fact recognize that grievance procedures constitute an effort to comply with civil rights law.

Event-history analyses of the diffusion patterns of EEO grievance procedures show that, consistent with institutional theory, attributes of the institutional environment rather than internal technical concerns drive the diffusion of EEO grievance procedures (compare Edelman 1990, 1992; Mezias 1990; Sutton et al. 1994). These analyses (shown in the appendix) reveal that, consistent with institutional theory, organizations with the highest rates of EEO grievance procedure creation were those closest to the public sphere (both government organizations and colleges are more than twice as likely as business organizations to create EEO grievance procedures). Furthermore, organizations with affirmative action offices were more likely to create EEO grievance procedures than those without, suggesting that these offices availed themselves of the professional literature recommending internal grievance procedures. Thus, the event-history analyses, which show a diffusion of grievance procedures, support our argument that there was a gradual institutionalization of the "rational myth" that internal grievance procedures would insulate organizations from their legal environments.

But how do we know that the "rational myths" circulated by the personnel profession, and apparently adopted by organizations, are not in fact "reality"? If EEO grievance procedures in fact produce efficiency benefits for organizations, then market-based theories would be largely correct. To answer this question, it is first worth noting that many of the claims we found in the personnel journals were based on generalizations from the experience of single personnel managers, or simply on suppositions.[7] As a result, it is impossible to trace the authority for these statements or to determine whether they are well-founded (see Galanter 1983, 1994).[8]

---

[7] As found in a study of the professional construction of wrongful discharge (Edelman et al 1992), it is quite common for such claims to lack any references to those studies or authorities backing up their conclusions. Consider the following 1985 article by an associate editor of *Personnel,* published by a division of the American Management Association: "According to recent surveys, the number of lawsuits initiated against firms by former or current employees [has] been steadily increasing over the past few years, along with the number of organizations that have instituted formal grievance procedures for nonunion workers. There are good reasons for organizations to institute these procedures: (1) to ensure fair treatment of employees across the board, (2) to raise morale in the workplace, and (3) to avoid costly court litigation" (Lo Bosco 1985, p. 61). Following the standard practices of journals of this type, this article provides no references to these "recent surveys," and we were unable to find any.

[8] In an article on discourse about lawyers, Galanter (1994, p. 664) calls attention to "a series of factoids or macro-anecdotes about litigation [that] became the received wisdom." Galanter (1983, p. 61) argues that knowledge "is not the mechanical recording of something out there—it is an interpretation of what we encounter, informed by our hopes and fears and by our pictures of how the world is."

Although we do not maintain that the claims of the professions are entirely wrong, we suggest that the professions' arguments often have little factual basis and are sometimes greatly exaggerated. We base our argument on an empirical examination of the two common claims in the personnel literature: First, that organizations with internal grievance procedures will experience fewer claims to external EEO agencies; and second, that courts will be less likely to find organizations in violation of civil rights law if they have an internal grievance procedure in place. The next two sections show that these claims were largely myth at the time they were formulated. In that sense, they provide evidence for institutional accounts of rationality. But we also show that over time certain economic benefits begin to accrue to organizations that adopt these procedures. Thus, EEO grievance procedures turn out to have some market-based rationality, although that rationality is socially constructed.

## THE MYTH OF GRIEVANCE PROCEDURES AS INSULATORS

As exemplified in the articles discussed above, a major appeal of EEO grievance procedures is their perceived capacity to internalize complaints, thus buffering the organization from complaints to external agencies and from lawsuits. To test this claim, we measure organizations' propensity to elicit both external and internal complaints, focusing in particular on the effects of EEO grievance procedures. We consider three ways in which the presence of EEO grievance procedures may affect the volume of discrimination complaints.

First, if special EEO grievance procedures reduce the likelihood of lawsuits, they must reduce the likelihood of complaints to external fair employment agencies, which employees must file as a first step in pursuing their statutory rights under Title VII and several other civil rights statutes. Thus, one measure of the efficacy of special EEO grievance procedures is their impact on the *number of external complaints*.

Second, the presence of an internal grievance procedure may in fact generate complaints that would not have been voiced were external complaint forums the only option. EEO grievance procedures may generate *internal* complaints by providing an alternative forum, which employees may see as less expensive, more accessible, or less likely to result in retaliation than external forums. Thus, the second effect of EEO grievance procedures we consider is the effect of those procedures on the overall *number of internal complaints*.

Third, EEO grievance procedures may also generate *external* complaints by making employees more aware of their rights or by encouraging internal complaints but then failing to provide satisfactory resolutions. This may motivate employees to file external complaints where they

would not have done so otherwise. To fully understand the dynamics of internal grievance procedures, then, it is necessary to assess the likelihood that complaints will be internalized, controlling for the overall volume of complaints. Thus, for our third measure of the efficacy of EEO grievance procedures, we estimate the effects of EEO grievance procedures on the ratio of internal to total complaints. This *"internalization ratio"* may represent the diversion of complaints from the external legal system. Alternatively, it may simply reflect an increase in different types of complaints, which employees are more likely to pursue in internal forums.

Some of the effects of a special EEO grievance procedure on the volume of internal and external complaints, and on the internalization ratio, may be influenced (or supplanted) by a general internal grievance procedure (i.e., one that is not specifically designated for discrimination-related complaints). Although we specified models estimating the effects of general grievance procedures as well, in this article we focus on special EEO grievance procedures because we expect that much of the impact of special EEO grievance procedures is due to their *symbolic* rather than their *functional* value. Special EEO grievance procedures tend to encourage internal vis-à-vis external complaints because they signal a commitment to fair treatment and to the resolution of discontent.

The volume of discrimination complaints is, of course, likely to be a function of a number of factors other than whether there is an EEO grievance procedure in place. Indeed, since employees are likely to be responsive to their employers' symbolic actions (Fuller 1993) and to their institutional environments, many of the factors that encourage responsiveness to legal norms among managers may also motivate complaints and lawsuits. Employees may be more likely to mobilize their rights in an organizational culture that encourages attention to law than in a highly repressive or discriminatory culture. Thus, we expect that factors that render organizations more sensitive to their legal environments (sector and contractor status), and the presence of internal offices and staff that demonstrate attention to civil rights law (EEO offices and the presence of an EEO counselor), will increase the overall volume of complaints, both external and internal. The presence of other internal structures is likely to be particularly important since these structures tend to generate agendas that foster attention to EEO ideals (Edelman et al. 1991; Edelman and Petterson 1999). EEO office staffs and counselors may develop commitments to EEO goals and encourage employees to challenge discrimination (Edelman et al. 1991). Further, EEO office staff and counselors are likely to encourage the use of internal procedures because they believe in their capacity to reduce external complaints and lawsuits. Finally, internal counselors increasingly see complaint resolution as a good management technique and may encourage the use of internal EEO grievance procedures

for that reason (Edelman, Erlanger, and Lande 1993). The percentage of female and minority employees in an organization may also affect the volume of complaints by altering the internal culture of the organizations. A greater proportion of minorities and women increases employers' opportunity for discrimination and may sensitize employees to their collective status and rights.

Another factor that would likely affect the volume of discrimination complaints is, of course, the actual amount of discrimination in the organization. We do not have a direct measure of discrimination, and given legal ambiguity as to what constitutes discrimination, it would indeed be quite difficult to find one. Prior discrimination lawsuits could serve as a proxy for discrimination. Just as complaints, however, prior lawsuits (which must originate as external complaints) are likely to be a product of the organization's legal environment and internal legal culture. In addition, organizations may change their behavior after being subject to a suit. Thus, although we examine the effect of prior discrimination lawsuits on complaints, we note that it is impossible to disentangle the possible explanations for the observed effects. Perhaps a better proxy for discrimination would be the sector and contractor status variables. Economists consistently note greater responsiveness to civil rights law (measured by the workforce shares of minorities and women) in the public sector and among federal contractors (Leonard 1984, 1986; Heckman and Wolpin 1976). If there is less actual discrimination closer to the public sphere, then one would expect fewer discrimination complaints in those sectors. However, if, as we suggest, the heightened responsiveness of organizations in these sectors creates a culture more conducive to rights-consciousness, then one would expect an increase in complaints in those sectors. Clearly, both factors may be at work, which could temper the observed differences in complaint volume among the sectors.

Although the factors above are of primary interest, we also consider the effects of other organizational attributes that may affect organizations' propensities to evoke complaints. We consider organizational size, since larger organizations will produce more complaints. Organizational age could affect complaints if organizational culture is imprinted at the time of the organization's birth. Professionalization may increase discrimination complaints both because of a greater awareness of legal rights and greater subjectivity in promotions and hiring. The presence of a personnel office, the number of attorneys who handle EEO matters, and the presence of a union could also increase awareness of legal rights. Several other environmental variables may also be relevant: in particular, we consider geographic region (South vs. other) and whether the organization is in a manufacturing or service sector.

The data for these analyses come from a mail follow-up survey con-

ducted as part of Edelman's 1989 EEO survey.[9] Approximately two-thirds, or 207, of the original 346 respondents returned the mail questionnaire. The mail survey asked for data on internal and external complaints for 1986. The year 1986 was chosen because it was sufficiently close to the survey date to permit accurate recall, yet sufficiently prior to the survey for organizations to know whether external complaints had been filed. If organizations did not have data for 1986 but did know the number of complaints for another recent year, they were asked to provide that information instead, and a few organizations provided information for a year before or after 1986.

These data clearly provide only rough indicators, especially with respect to internal complaints, where organizations may have different criteria for recording a complaint. Further, any data on the timing of complaints may be problematic because it is difficult to know when complaints, external or internal, will affect organizations. Employers and personnel professionals may be aware of a developing case, for example, long before any complaint is filed. Or knowledge of one complaint may lead to more informal complaints at lower levels of an organization, which might not be officially recorded. However, the data on the relative frequency of internal and external complaints are, to our knowledge, unique and provide an important indicator of a phenomenon that is both theoretically and socially important.[10]

Because the theoretical construct we measure is organizations' propensity to evoke complaints, we estimate maximum likelihood Tobit models. Tobit analysis is appropriate where the dependent variable is censored at some upper or lower bound as a result of the way the data are collected (Winship and Mare 1992; Tobin 1958; Maddala 1983; Roncek 1992). In our case, we have censoring at a lower bound: an organization's propensity to elicit complaints could be less than zero since an organization, given

---

[9] The survey, Organizational Response to EEO/AA Law, was supported by a grant from the National Science Foundation (SES 88–14070) and is described in appendix A of Edelman (1992).

[10] Five variables were missing information for 0.5%–7% of the sample. We created missing data indicator variables (coded as "1" if missing, "0" if observed) for each of these variables and recoded the original indicators to the variable mean (Little and Rubin 1987). None of the missing value indicator variables are statistically significant in any of the models presented below, suggesting that organizations missing information on these measures do not have different rates of complaints than organizations with complete data. If the organization knew that it had a specific procedure, but did not know the year it was created, we assume that the procedure was in place in 1986. To test this assumption, we estimated models that included a dummy variable coded "1" if the year was unknown and "0" if known. This coefficient was also not significant and close to zero in magnitude.

its environment, structure, and internal legal culture, may discourage complaints as well as produce or encourage complaints. Since a majority of the organizations experienced no complaints in the survey year, the distribution of the dependent variables is skewed and least squares regression would produce biased and inconsistent estimates. We, of course, only observe the portion of the distribution that is zero or greater.[11] Thus, we estimate models of the form:

1. $Y * 1_i = X_iB + e_i$,
2. $Y1_i = Y * 1_i$ if $Y * 1_i > 0$,
3. $Y1_i = 0$ if $Y * 1_i \leq 0$,

where, for the $i$th observation, $Y * 1_i$ represents an organization's propensity to evoke complaints (the unobserved continuous latent variable), $Y1_i$ represents observed complaints; $X_i$ is a vector of values on the independent variables, $e_i$ is the error, and B is a vector of coefficients. We assume that $e_i$ is uncorrelated with $X_i$ and is independently and identically distributed (Winship and Mare 1992; Roncek 1992). Below we discuss Tobit models predicting the volume of complaints to external agencies, internal complaints, and the internalization ratio, in turn. Tobit coefficients are difficult to interpret because they capture two types of effects in a single estimate. We use Roncek's method (1992; see also McDonald and Moffitt 1980; Caspi et al. 1998) to partition these estimates into a component for the change in the probability of experiencing a complaint (among cases with zero complaints) and a component predicting the number of complaints (among organizations with at least one complaint). Table 2 shows the variables used in our analyses as well as descriptive statistics for those variables.

The most direct measure of the buffering capacity of EEO grievance procedures is their relationship to the volume of complaints to external fair employment agencies. Table 3 shows a Tobit model predicting the volume of complaints to these external agencies. To provide the strongest test of the buffering hypothesis posited by the legal and personnel professions, we construct the EEO procedure indicator as a dummy variable, coded "1" for organizations that have a specific EEO grievance procedure and "0" for organizations that have either no procedure or a general griev-

---

[11] The variable could also be conceptualized as a simple count, making Poisson regression the appropriate analysis technique. For all of our analyses, Poisson regression produces similar results, albeit with coefficients and $t$-values of greater magnitude. We think the Tobit results are more realistic, and they are in any case more conservative. OLS regressions also produced similar results.

TABLE 2

COMPLAINT ANALYSIS: VARIABLES, DESCRIPTIVE STATISTICS, AND VARIABLE
DEFINITIONS

| Variables | Mail Mean | Phone Mean | Description |
|---|---|---|---|
| Fixed independent variables:* | | | |
| Business | .71 (.45) | 69 (.47) | Private firm |
| College | .14 (.35) | .15 (.36) | College or university |
| Government | .14 (.35) | .16 ( 37) | Government agency |
| Size | 7.03 (15.2) | | Employees in hundreds (1984) |
| Log size | 5.57 (1.52) | 5.58 (1.54) | Natural log of full-time employees (1989) |
| %female | .46 (.25) | | Percentage of full-time permanent employees that are women (1984) |
| %minority | .17 (.19) | | Percentage of full-time employees that are members of racial or ethnic minority groups |
| %salary | .42 (.28) | .41 (.28) | Percentage of full-time employees that are salaried |
| South | .32 (.47) | .32 (.47) | "1" if South, "0" if other |
| Changing independent variables:* | | | |
| Contractor | .35 (.48) | .38 (.49) | Federal contractor subject to OFCCP regulation |
| EEO office | .17 (.38) | .17 (.38) | A separate department for equal employment opportunity or affirmative action |
| Personnel | .73 (.44) | .72 (.45) | Personnel department |
| Union | 41 (.49) | .44 (.50) | Partially or fully unionized |
| Lawsuit | 26 (.44) | .27 (.44) | Past suit for discrimination-related complaint |
| Fixed dependent variables:* | | | |
| Internal† | 1.56 (3.25) | | Number of EEO-related complaints filed through the organization's internal complaint procedure (1986) |
| External† | 1.41 (3.09) | | Number of EEO-related complaints filed with the EEOC or state agency (1986) |
| Internalization | .51 (.36) | | Ratio of internal to total complaints (1986) |
| Changing dependent/independent variable:* | | | |
| EEO procedure | .27 (.44) | .24 (.43) | Specific equal employment opportunity grievance procedure |

NOTE —SDs are given in parentheses. $N$ = 206 for mail responses and 345 for phone responses.

* Fixed variables retain the same value in each period or are available for only one period. Changing variables were collected as event histories and are coded "0" before the event occurs and "1" after the event occurs

† Four outliers were top-coded at 15 complaints. The internalization ratio is constructed from the original, rather than the top-coded, complaint data Two organizations that processed complaints for other agencies were dropped from the analyses

TABLE 3

MAXIMUM LIKELIHOOD TOBIT REGRESSION OF EXTERNAL COMPLAINTS

| Variable | Coefficient | T-ratio | Δ in Number | Δ in Probability (%) |
|---|---|---|---|---|
| Model 1: | | | | |
| Constant .... ................ .. ......... | −4.57*** | −3.75 | | |
| EEO procedure  ........ ......... .. . | .62 | .63 | .19 | 5 |
| Size (hundreds) ...... ... .............. | .17*** | 6.85 | .05 | 1 |
| %female ....... ... .......... .. ...... | 1.04 | .59 | .32 | 8 |
| %minority ....... .. .......... . ....... | 3.72* | 1.68 | 1.13 | 29 |
| Government sector .. ......... .. | .15 | 12 | .05 | 1 |
| College sector . .......... . .. ........ | −.25 | −.19 | −.08 | −2 |
| Contractor (business) ... .. ...... | 1.92* | 1.82 | .58 | 15 |
| Sigma ............... . ......... ... ..... | 4 87*** | 11.40 | | |
| Log-likelihood ..... .... ............. | −294.66 | | | |
| N .............. .............. ... .............. | 205 | | | |
| Model 2: | | | | |
| Constant ........... .. ........... .. .. | −6.21*** | −4.27 | | |
| EEO procedure ... .... ......... .... | −.44 | −.46 | −.13 | 4 |
| Size (hundreds) ......... . .. ........ | .12*** | 4.95 | .04 | 1 |
| %female .......... ......... .. .. ...... | 1 19 | .71 | .36 | 10 |
| %minority .......... .. .............. ... | 2 64 | 1.24 | .80 | 22 |
| Government sector ............... .. | −1.05 | −.83 | −.32 | −9 |
| College sector ............ . .. ........ | −1.50 | −1 19 | −.46 | −13 |
| Contractor (business) .. .. .. ... | 88 | .87 | 27 | 7 |
| EEO office ............ .... ........ .. . | 3.10*** | 2.86 | .94 | 26 |
| EEO counselor ......... .............. | 2.40** | 2.03 | .73 | 10 |
| Lawsuit ........ ............... ............ | 2.29** | 2.48 | .69 | 18 |
| Sigma .. ............. ......... ... ... ..... | 4 51*** | 11.51 | | |
| Log-likelihood ..... ............... .. .. | −284.52 | | | |
| N .............. .............. .... ........ .. . | 205 | | | |

* P < .10.
** P < 05
*** P < .01.

ance procedure. Model 1 shows EEO procedure and organizational and sectoral control variables. In model 2, we add other EEO characteristics that may affect an organization's propensity to elicit complaints.

Table 3 shows that specific EEO grievance procedures have little effect on the volume of complaints to external fair employment agencies. In model 1, the coefficient is close to zero and positive, and in model 2 the coefficient is close to zero and negative. This is the strongest evidence that managers' common beliefs in the buffering capacity of such procedures are probably wrong. Without the additional EEO variables (model 1),

only organizational size,[12] the percentage minority, and government con-
tractor status have even marginally statistically significant effects on the
volume of external complaints.[13] Approximately 30% of the total effect of
each independent variable is on increasing or decreasing the number of
complaints among organizations that had at least one complaint, and
about 70% of the effect is on increasing or decreasing the probability of
experiencing a complaint among organizations that had no complaints.[14]
Among organizations with complaints, a 1,000-worker increase in full-
time employees corresponds to an additional one-half of one complaint;
among organizations without complaints, it corresponds to a 10% higher
probability of experiencing a complaint.[15] A 10% increase in the minority
workforce similarly corresponds to about one-tenth of one complaint
among organizations with complaints and raises the probability of experi-
encing a complaint by about 3%. Businesses with government contracts
were about 15% more likely to have a complaint and experienced about
0.6 more complaints than noncontractor businesses.

Including the additional EEO characteristics in model 2 significantly
improves the fit of the model and mediates the effects of the percentage
of minority employees and the contractor status. As expected, organiza-
tions with EEO offices and EEO counselors have a greater propensity to
elicit complaints since offices tend to call attention to civil rights and staff

---

[12] We also estimated a series of models predicting the *ratio* of complaints to organiza-
tional size. Our substantive findings are robust under this specification: EEO proce-
dures again fail to significantly alter the level of complaints to external agencies.

[13] Note that the sector variables (government and college) are not generally statistically
significant in these analyses. These are key variables in institutional analyses and
are generally found to be statistically significant in modeling diffusion processes (e.g.,
Edelman 1990, 1992). Their lack of significance here suggests that vulnerability to
the institutional environment does not directly drive the complaint process.

[14] We obtain these numbers using Roncek's (1992) equation 4a for organizations with
at least one external complaint, $B_1 \times [1 - (z \times f(z)/F(z)) - f(z)^2/F(z)^2]$, where $B_1$ is
the Tobit coefficient for a particular independent variable, $f(z)$ is the unit normal
density or the value of the derivative of the normal curve at z, $F(z)$ is the cumulative
normal distribution function for the proportion of cases with at least one complaint,
z is the z-score associated with the area under the normal curve, and sigma is the
standard deviation of the error term in the estimated equation. For model 1 of table
3, $B_1 \times \{1 - (-.3042).3807/.3805\} - (.3807^2/.3805^2) = B_1 \times .3033$. For organizations
not experiencing complaints, the change in the cumulative probability of having a
complaint is $B_1 \times f(z)/sigma$ or $B_1 \times .0782$ (from Roncek 1992, p. 504, eq. 4b).

[15] We obtain these interpretations for organizations with complaints and without as
follows:
for size, $B = .17 \times .3033 = .05$ per hundred (or .5/thousand) and $.17 \times .0782 = .013/$
hundred or 13%/thousand; for percentage minority, $B = 3.72 \times .3033 = 1.13$ ($\times .1$)
$= .11$ and $3.72 \times .0782 = .29$ ($\times .1$) = 3% per 10% increase; for contractor, $B = 1.92$
$\times 3033 = .58$ and $1.92 \times .0782 = 15$.

may encourage employees to pursue those rights (Edelman et al. 1991). Having an EEO office increases the expected number of complaints by almost one full complaint (3.1 × .3033 = .94) and raises the probability of experiencing any complaint by 26% (3.1 × .0844 = .26). Having experienced a previous discrimination-related lawsuit is also a strong predictor of the volume of external complaints, increasing the expected number of complaints by 0.7 and raising the probability of experiencing any complaint by 18%. As discussed above, this probably represents a combination of the effect of complaints motivated by discrimination and the heightened rights awareness that may accompany lawsuits.

A number of factors, found to have no statistically significant effect, are omitted from the estimated models shown in table 3. Other than organizational size and contractor status, which have the expected positive effects on external complaints,[16] organizational characteristics (age, professionalization, and the presence of a union) have small nonsignificant effects. Although preliminary analyses had shown the number of lawyers employed by the organization to have a positive effect on the number of external complaints, we omit this factor because we are unsure of the temporal ordering of attorneys and complaints. Variables representing region, core/periphery, and manufacturing/service are not statistically significant.

In addition to their apparent inability to buffer organizations from *external* complaints, table 4 shows that EEO grievance procedures may increase the volume of *internal* discrimination complaints. The effect of EEO grievance procedures on internal complaints is statistically significant in model 1, but this effect is mediated by EEO offices, EEO counselors, and lawsuits in model 2. In model 1, having an EEO procedure increases the expected number of complaints by more than one-half of one complaint (2.15 × .2976 = .64) and raises the probability of experiencing a complaint by 18% (2.15 × .0821 = .177). EEO counselors are especially likely to channel employees toward internal forums, increasing the expected number of complaints in model 2 by 1.4 and raising the probability of experiencing a complaint by 42%. Since EEO counselors often handle discrimination complaints and problems, they are likely to be vested in the internal resolution of complaints (Edelman et al. 1993). As with external complaints, there is a marginally higher volume of internal complaints in

---

[16] In organizational analysis, it is common to use the log of size rather than raw size because it is generally the case that an increase in the size of a relatively small organization will have a greater effect on organizational structure than an increase in the size of larger organization. In predicting the number of complaints, however, we do not log size since every additional employee would appear to add an equal chance of a discrimination complaint. We do represent the size variable in hundreds of employees, however, so that the coefficients provide more information.

TABLE 4

MAXIMUM LIKELIHOOD TOBIT REGRESSION OF INTERNAL COMPLAINTS

| Variable | Coefficient | t-ratio | Δ in Number | Δ in Probability (%) |
|---|---|---|---|---|
| Model 1. | | | | |
| Constant ................................ ... | −4 66*** | −4.03 | | |
| EEO procedure ....................... | 2.15** | 2.35 | .64 | 18 |
| Size (hundreds) . .. .. ................. | .19*** | 8.35 | .06 | 2 |
| %female .. ... .... .. .. .. ... .. .. | −.69 | −.41 | −.21 | −6 |
| %minority ............................... | 5.55*** | 2.65 | 1.65 | 46 |
| Government sector ............ .. | .82 | .69 | .24 | 7 |
| College sector .. . ... .............. | 1.95* | 1.68 | .58 | 16 |
| Contractor (business) . . .. .. .. | 1.07 | 1.05 | 32 | 9 |
| Sigma ....................... .. ..... . . .. | 4.60*** | 11.09 | | |
| Log-likelihood ............ ....... . ... | −284.99 | | | |
| N .. .... .. .. .. ... . .. ............ ... | 205 | | | |
| Model 2 | | | | |
| Constant .. .... .. . .. ..... .. . ... | −8 15*** | −4.88 | | |
| EEO procedure .............. ... ..... | 1 28 | 1.47 | .38 | 12 |
| Size (hundreds) ...... .. .. ............ | .15*** | 6.57 | .05 | 1 |
| %female .. ... .. . ... .. ... .. .. | 1.16 | 1.03 | .35 | 10 |
| %minority .... .. .. ..... ... .. ... | 4.88** | 2.34 | 1.45 | 44 |
| Government sector ............ . ... | .07 | .05 | .02 | 1 |
| College sector ... .. ...................... | 1.16 | 1.03 | .35 | 10 |
| Contractor (business) .. .. .. .. | .18 | .19 | .06 | 2 |
| EEO office ..... ... .. .. ..... .. ... | 2.01** | 2.00 | .60 | 18 |
| EEO counselor .................. .. ... | 4.69*** | 3.38 | 1.40 | 42 |
| Lawsuit . ....................... ............. | 1.64* | 1.88 | .49 | 15 |
| Sigma ... .. ..... ........... .......... | 4.25*** | 11.24 | | |
| Log-likelihood . . .. ... .. .... | −272.22 | | | |
| N . .. .. ..... .... . . .... .. ... .. ... | 205 | | | |

* P < .10.
** P < 05.
*** P < .01.

organizations that have experienced a prior discrimination-related law-suit, suggesting effects of either greater discrimination or greater rights awareness. As in the analysis of external complaints, only size and the percentage of minority employees had a statistically significant effect among the other organization-level or sector-level variables.

The results presented so far, then, suggest that EEO grievance proce-dures have relatively small effects on the volume of external and internal complaints. However, their presence indirectly affects the volume of inter-nal complaints because EEO offices and EEO counselors appear to en-

courage their use. This raises the possibility that, controlling for the volume of total EEO complaints, specific EEO grievance procedures may be associated with greater internalization of these complaints. By modeling the *internalization ratio,* that is, the ratio of internal complaints to total (internal and external) complaints, we determine the effect of EEO grievance procedures and other organizational and environmental factors on the internalization of complaints.

Because organizations must have experienced at least one complaint to calculate the internalization ratio, there is a possibility of selectivity bias in estimating the ratio; factors that affect whether organizations experience any complaints are likely to be associated with internalization. To control for this possibility, we use Heckman's (1976) selectivity approach to model jointly the process generating any complaint and the process generating the internalization of complaints.[17] This is a generalization of the Tobit model, in which $Y * 2_i$, representing whether any complaints are filed, affects whether $Y1_i$ (the internalization ratio) is observed (see Winship and Mare 1992). Thus, this model takes the form:

1. $Y * 1_i = X_iB + e_{1i}$,
2. $Y1_i = Y * 1_i$ if $Y * 2_i > 0$,
3. $Y1_i = 0$ if $Y * 2_i \le 0$.

Table 5 shows the selection model predicting whether organizations will have at least one complaint. This equation produces a coefficient, lambda, that we use to statistically adjust the internalization ratio equation for selectivity. Heckman's method is most appropriate when there is a determinant of the selection equation that can be excluded from the substantive equation. One such exclusion is unionization: unionized organizations are more likely than nonunionized organizations to experience complaints, but there is less reason to expect unionization to affect internalization.[18] Table 5 is also substantively relevant, showing that specific

---

[17] For this part of the article, we would prefer to have information at the level of the individual complaint rather than at the organizational level of analysis. Although we can recover the proportion of internal to total complaints for each organization, data on the type and fate of individual complaints would speak more directly to the efficacy of EEO grievance procedures and processes of internalization. In addition to the internalization ratio selection model discussed below, we also estimated nonratio models predicting internal and external complaints using total complaints as a regressor. We also specified ratio models with 1/total complaints (Gibbs and Firebaugh 1985, p. 717) as a predictor and censored regression models with an upper limit of one and a lower limit of zero. In each case, the substantive results with respect to EEO procedures and counselors parallel those presented in table 6.

[18] Since no exclusionary restriction is ideal, we also specified models omitting factors such as organizational size, region, and industry from the substantive equation. When

429

TABLE 5

SELECTION EQUATION PREDICTING ANY COMPLAINT

| | MODEL 1 | | MODEL 2 | |
| VARIABLE | Coefficient | t-ratio | Coefficient | t-ratio |
|---|---|---|---|---|
| Constant ............ ...... ..... .. ......... | −.93*** | −3.44 | −1.50*** | −4.38 |
| EEO procedure ................. .. ...... | .66*** | 2.67 | .44* | 1.65 |
| Size (hundreds) ... ...................... ..... | .06*** | 3.87 | .04*** | 2 69 |
| %female ..................... ..... ........... | .09 | .23 | .06 | .14 |
| %minority ...... ................... .. .. ........ | .65 | 1.29 | .50 | .95 |
| Government sector .... ................... | −.33 | −1.10 | −.51 | −1.53 |
| College sector ........... ..................... | .23 | .29 | .07 | .23 |
| Contractor......... ................... ..... .... | .78*** | 2.98 | .66** | 2.42 |
| Union (business) ... ........................... | .27 | 1.35 | .29 | 1.38 |
| EEO office ............. ..... .................. | | | .50 | 1 56 |
| EEO counselor ............ .. ................ | | | .83*** | 3.08 |
| Lawsuit ............ ............................... | | | .28 | 1.39 |
| Log-likelihood ................................ | −112.66 | | −104.87 | |
| N ... .. .................. ...... .................. ... | 205 | | 205 | |

* P < .10
** P < .05
*** P < 01.

EEO grievance procedures have a statistically significant effect on the likelihood that organizations will experience at least one discrimination complaint. This is notable because many organizations that do not have special EEO grievance procedures already have general grievance procedures in place that could be used for discrimination complaints. Federal contractors and organizations with EEO counselors are also more likely to have at least one complaint. Not surprisingly, size is statistically significant. Table 6 shows models predicting the internalization ratio, adjusted for selectivity bias. EEO grievance procedures have a marginally statistically significant effect on internalization, standardized by overall volume. Colleges, relative to private businesses, are particularly successful at internalizing EEO complaints. EEO counselors also have a strong effect on the internalization ratio. There are several possible interpretations of the effect of special EEO grievance procedures on the internalization ratio. First, if these procedures exert a small negative effect on the volume

---

these models (which include unionization in both equations) are tested, the coefficient for unionization in the internalization ratio equation is close to zero (-0.05) with $p$-values of approximately 0.5. Under these specifications, however, the models are less stable and the overall fit is worsened. The zero-order correlation is -0.008 between unionization and internalization and 0.16 between unionization and any complaint.

TABLE 6

SELECTION EQUATION PREDICTING INTERNALIZATION RATIO

| | MODEL 1 | | MODEL 2 | |
| --- | --- | --- | --- | --- |
| VARIABLE | Coefficient | t-ratio | Coefficient | t-ratio |
| Constant . ......................................... | .28 | 1.10 | −.22 | −.48 |
| EEO procedure ..... ...... ...... ... ... . . | .21* | 1.93 | .19* | 1.79 |
| Size (hundreds) .. ... . ......... .. ........ | .004 | 1.04 | .004 | 1.25 |
| %female .......... .. ............................. | −.20 | −1.20 | −.28* | −1.67 |
| %minority ................................ ... ... | 18 | .82 | .14 | 62 |
| Government sector ..... .. _ .. ..... .. | .002 | .02 | −.02 | −.17 |
| College sector .. .... ...... ... _ ...... ...... | .17 | 1.54 | .20* | 1.67 |
| Contractor (business) ..................... | .06 | .53 | .06 | .53 |
| EEO office ................................. ...... | | | .04 | .35 |
| EEO counselor ....................... .. ..... . | | | .52** | 2.35 |
| Lawsuit ... .. ..................................... | | | −.02 | −.29 |
| Selection correction ................... .. | .21 | 1.00 | .33 | 1.23 |
| Log-likelihood ................. ... ..... .. | −28.49 | | −23 85 | |
| N .................. ...... .. .. ....... ..... ......... | 205, 98 | | 205, 99 | |
| R² ........ .. ....................................... | .10 | | .15 | |

* P < .10.
** P < .05.
*** P < 01

of external complaints and a small positive effect on the volume of internal complaints, the difference could produce a statistically significant effect on the internalization ratio. Second, if EEO offices and counselors drive up the volume of total complaints, special EEO grievance procedures could divert a portion of the excess volume to internal forums. Third, and we think this is the most likely explanation, the presence of an EEO grievance procedure together with the actions of EEO offices and EEO counselors may motivate different types of complaints, which are more likely to be pursued in internal forums. In particular, whereas external forums are most likely to be used for claims involving discrimination in hiring and firing (Donahue and Siegelmann 1991), internal grievance procedures may be more likely to attract claims involving working conditions and, in particular, sexual harassment. The positive (though nonsignificant) value for the selection correction term suggests that organizations experiencing at least one complaint may possess unmeasured characteristics, such as rights consciousness, for example, that are positively related to internalization.

Taken together, the analyses summarized in tables 3–6 show that special EEO grievance procedures generally fail to insulate organizations from the threat of complaints to external fair-employment agencies. Addi-

tional analyses (not shown) suggest that general grievance procedures also had no insulating effect. Thus, many of the claims of the personnel, legal, and management professions about the rationality of these procedures are largely mythical. The presence of EEO grievance procedures is associated with a greater propensity to elicit internal complaints, although it appears to be EEO offices and EEO counselors, rather than the procedures themselves, that drive the volume (and possibly the nature) of these complaints. Thus, EEO grievance procedures may play an important role in resolving complaints within the firm, but they do not appear to protect organizations from experiencing external complaints or litigation.

## THE MYTH OF JUDICIAL DEFERRAL TO GRIEVANCE PROCEDURES

In this section, we examine the claims of the personnel and legal professions that courts will look favorably upon organizations with grievance procedures. In particular, we examine evidence regarding the two major assertions that we found in the personnel literature: (1) employers with internal grievance procedures in place are less likely to be found in violation of civil rights law because the grievance procedures will serve as evidence of fair treatment; and (2) if an employer has an internal grievance procedure, courts may dismiss an employee's lawsuit for failure to use that grievance procedure.

These arguments had little legal foundation at the time that the personnel literature began to articulate them since virtually all discrimination claims were at that time decided under a *vicarious liability standard.* Under that standard, employers are held responsible for the wrongful acts of their employees regardless of whether they knew about the wrong-doing. Thus, neither a policy against discrimination nor a grievance procedure would help an employer escape liability. Not surprisingly, few cases prior to the mid-1980s even discussed the relevance of grievance procedures.

In contrast to the assertions of articles such as those quoted earlier, the vicarious liability standard applied to sexual (and also racial) harassment cases as well as to discrimination cases through the mid-1980s. Since the 1964 Civil Rights Act did not explicitly prohibit sexual harassment, courts had to decide whether it constituted sexual discrimination under the law. Prior to 1986, courts were most responsive to allegations of quid pro quo sexual harassment, in which supervisors demanded sexual favors as a condition of employment or retaliated for denial of those favors. The judicial position on the role of grievance procedures under the vicarious liability standard is apparent, for example, in the federal court of appeals case of

*Miller v. Bank of America* (600 F.2d 211 [1979]).[19] In *Miller v. Bank of America,* the plaintiff sued because she was fired after refusing her supervisor's demand for sexual favors from a "black chick." Bank of America argued that it should not be liable because it had an established policy against harassment and had afforded Ms. Miller an opportunity for redress through its internal grievance procedure. Citing similar cases from four other federal appellate circuits, the court held that, under the doctrine of respondeat superior (a type of vicarious liability), employers are liable for the acts of their agents (here, the supervisor) regardless of any policies against those acts and regardless of the employer's knowledge of those acts. It held, moreover, that Title VII does not require employees to use internal grievance procedures before filing formal legal claims.[20]

---

[19] There was some variation in the courts on the issue of vicarious liability at this time. One federal district court held that an employer could not be held liable for the sexual harassment of an employee by a supervisor unless there was an explicit company policy condoning the harassment (*Corne v. Bausch & Lomb,* 390 F.Supp. 161 [1975]). Another specifically rejected the logic in the Corne case and held that sexual harassment by a supervisor was a violation of Title VII as a matter of law because it created an artificial barrier to employment that burdened one gender but not the other (*Williams v. Saxbe,* 413 F.Supp. 654 [1976]). A third case treated the supervisor as an agent of the employer, which meant that a policy or practice of a supervisor constituted a policy or practice of the company under the doctrine of respondeat superior (*Munford v. Barnes,* 441 F.Supp. 459 [1977]). In Munford, the court arguably raised the possibility that an employer could take some affirmative action that would insulate it from liability by suggesting that the employer was responsible because it had not investigated the harassment even after it was given notice. But since then, courts generally have a vicarious liability theory in quid pro quo cases, which means that the employer is responsible whether or not the employer knew, should have known, or approved of the supervisor's actions (e.g., *Anderson v. Methodist Evangelical Hospital, Inc.,* 464 F.2d 723, 725 [CA6 1972]).

[20] Interestingly, the Linenberger and Keaveny article (1981) mentioned earlier cites the Miller case in suggesting that grievance procedures would be valuable in litigation. This may be based on a comment at the end of the case suggesting that the holding will not put an undue burden on employers because, if the EEOC determines that there is reasonable cause to believe that discrimination has occurred, it must allow for conciliation. The court suggests that "an employer whose internal procedures would have redressed the alleged discrimination can avoid litigation by employing those procedures to remedy the discrimination upon receiving notice of the complaint or during the conciliation period" (*Miller v. Bank of America,* 214). This does not suggest, however, that if the case does eventually go to court, the court will be impressed by the grievance procedure. The Linenberger and Keaveny article also refers to the EEOC guidelines in making its claim. Here, ambiguity in the guidelines may have contributed to the article's claims. The first guidelines on sexual harassment, issued in 1980, state explicitly that employers are vicariously liable for the acts of their agents. "Applying general Title VII principles, an employer . . . is responsible for its acts and those of its agents and supervisory employees with respect to sexual harassment regardless of whether the specific acts complained of were authorized or even forbidden by the employer and regardless of whether the employer knew or

A new theory of sexual harassment—the *hostile work environment* theory—began to evolve in the early 1980s (Hipp 1988). Under this theory, an employee may sue because coworkers' harassing acts make the workplace intolerable for the plaintiff. Because this new type of sexual harassment often involves harassment by coworkers rather than by supervisors, the issue of agency is more problematic and thus the courts look for signs of ratification of the harassment by the employer rather than simply holding the employer vicariously liable.

The hostile work environment theory was formally recognized by the United States Supreme Court in the 1986 case of *Meritor Savings Bank v. Vinson* (106 Sup. Ct. 2399 [1986]). The plaintiff, Michelle Vinson, alleged that her supervisor, Sidney Taylor, made repeated demands for sexual relations, fondled her at work, and raped her on several occasions. Vinson testified that she did not report Taylor's behavior or use the company's grievance procedure because of her fear of Taylor. Meritor Savings Bank argued that because it did not know of any sexual misconduct by Taylor, it could not be held responsible. In its review of the case, the court of appeals used a vicarious liability standard and held that, irrespective of the availability of the grievance procedure, the bank was liable for Taylor's conduct. The Supreme Court then reversed the court of appeals, recognizing the hostile environment cause of action and replacing the vicarious liability standard with a standard of *direct liability* in which only employers who knew or should have known about the discriminatory acts

---

should have known of their occurrence" 29 C.F.R. §1604.11(c). However, the next section is ambiguous and seemingly contradictory and suggests that an employer's subsequent actions may sometimes affect liability. "With respect to persons other than those mentioned in paragraph (c) of this section, an employer is responsible for acts of sexual harassment in the workplace where the employer, or its agents or supervisory employees, knows or should have known of the conduct. An employer may rebut apparent liability for such acts by showing that it took immediate and appropriate corrective action" 29 C.F.R. §1604.11(d). One year later, paragraph (d) was revised to state "With respect to conduct between fellow employees." Thus, vicarious liability would seem to apply only to supervisors and not to fellow employees. Finally, paragraph (e) raises the issue of grievance resolution, although it suggests that employers should inform employees of their right to complain to the EEOC and does not mention internal complaints. "Prevention is the best tool for the elimination of sexual harassment. An employer should take all steps necessary to prevent sexual harassment from occurring, such as affirmatively raising the subject, expressing strong disapproval, developing appropriate sanctions, informing employees of their right to raise and how to raise the issue of harassment under Title VII, and developing methods to sensitize all concerned" 29 C.F.R §1604.11(e). The suggestion in the EEOC guidelines that agency principles might sometimes apply, together with paragraph (e), which raises the issue of employers' responsibilities, may have led Linenberger and Keaveny (1981) and others to recommend internal grievance procedures. But nothing in the guidelines, or any cases we were able to find, specifically suggests that employers may avoid liability by creating internal grievance procedures.

of their agents are liable. With the adoption of the direct liability standard, the earlier claims of personnel professionals suddenly gained validity since employers could argue that their grievance procedures constituted evidence that they sought to learn about and correct their agents' misdeeds. Though the court in the Meritor case held that the grievance procedure in question was inadequate to insulate the employer from liability (because it required the victim to complain directly to her alleged harasser), it noted that in a future case, a better grievance procedure might provide such insulation. "The bank's grievance procedure apparently required an employee to complain first to her supervisor, in this case Taylor. Since Taylor was the alleged perpetrator, it is not altogether surprising that respondent failed to invoke the procedure and report her grievance to him. Petitioner's contention that respondent's failure should insulate it from liability might be substantially stronger if its procedures were better calculated to encourage victims of harassment to come forward" (*Meritor Savings Bank v. Vinson*, 72–73).

Thus, although the court found the grievance procedure in question in the Meritor case inadequate, it reinforced the personnel professions' claims that grievance procedures may insulate employers from liability. Almost immediately after that case was decided, a federal circuit court of appeals adopted a similar standard in racial harassment cases (*Hunter v. Allis-Chalmers*, 797 F.2d 1417 [1986]). Racial harassment cases are quite similar to hostile work environment sexual harassment cases in that they involve acts that may make the workplace intolerable for minorities but do not necessarily involve direct threats to one's economic welfare.

Further, a recent Supreme Court decision—*Faragher v. City of Boca Raton* (118 Sup. Ct. 1115 [1998])—both reinforces the decision in *Meritor Savings Bank v. Vinson* on hostile environment sexual harassment and gives credence to the personnel profession's claim that employers may avoid liability where they have an internal grievance procedure in place and an employee fails to use it.[21] The court held, "While proof that an employee failed to fulfill the . . . obligation of reasonable care to avoid

---

[21] A concurring opinion in the Meritor case also suggested that an employee's failure to use an internal complaint procedure might render a court reluctant to find that the employee was constructively discharged, since the employee had waived an opportunity to redress her grievance. The concurring opinion, written by Justice Marshall and joined by Justices Brennan, Blackmun, and Stevens, would have retained the vicarious liability standard but nonetheless suggests that an employer's attempts to redress grievances may be relevant to the issue of remedies. In this regard, the opinion states that "where a complainant without good reason bypassed an internal complaint procedure she knew to be effective, a court may be reluctant to find constructive termination and thus to award reinstatement or backpay" (*Meritor Savings Bank v. Vinson*, 2399, 2411).

harm is not limited to showing an unreasonable failure to use any complaint procedure provided by the employer, a demonstration of such failure will normally suffice to satisfy the employer's burden [of proof]."

The law on hostile environment sexual harassment cases since 1986 (and to an even greater extent in 1998), then, very much fits with the claims made by the personnel profession in the late 1970s and early 1980s. But the law on virtually all other types of discrimination claims remains one of vicarious liability: an employer is responsible for the acts of its agents irrespective of an internal grievance procedure. Even the law on hostile environment sexual harassment claims, moreover, developed *after* the personnel professions' claims regarding the value of internal grievance procedures, as they were presented in the late 1970s and early 1980s. It would appear, then, that judicial recognition of grievance procedures did not motivate personnel professionals' claims and organizations' creation of grievance procedures but rather that *the courts were following institutionalized organizational practices.* The only other explanation, which we examine below, would be that the court in the Meritor case was recognizing a trend that had become well established in the lower courts. This does not, however, appear to be the case.

To determine when the law began to recognize grievance procedures as a source of protection for employers and the extent to which internal grievance procedures insulate employers from legal liability, we conducted a content analysis of all available federal Title VII cases involving sexual or racial discrimination or harassment in which employers discussed "the grievance procedure defense"; that is, they asserted that they should not be held liable for discrimination because they had a grievance procedure in place.[22]

We searched the Westlaw federal case database for all cases from 1964 through 1997, which alleged sexual or racial discrimination or harassment under Title VII of the 1964 Civil Rights Act and also addressed the relevance of grievance procedures.[23] The search found a total of 477 cases, none prior to 1970. Of these, 161 cases were excluded for one or more of the following reasons: they were not Title VII cases (or the part of the case addressing grievance procedures did not involve Title VII); they did not in fact involve the grievance procedure defense; they did not involve racial or sexual discrimination or harassment; or they were duplicate cases. We also eliminated 177 cases that involved union grievance proce-

[22] In most cases, the grievance procedure defense is asserted in combination with other defenses against liability.

[23] The search request, which was conducted for each year within the period 1964–98 within the federal case database, was "Title VII" & race racial sex! /P "grievance procedure!".

dures, which tend to be decided by reference to collective bargaining agreements and therefore cloud the issue of judicial deferral to organizational grievance procedures. This left us with 129 cases.

Westlaw is one of the two major on-line databases that are used by the legal profession to discover precedent (the other is Lexis).[24] Since attorneys tend to rely heavily on on-line case reporters, these databases are the de facto case law that influences judges, lawyers, and employers.[25] Because we only searched for cases that discussed grievance procedures, the data do not speak to the question of what proportion of all discrimination cases involve an employer's claim that the grievance procedure should provide insulation. The cases do tell us, however, when the grievance procedure defense first arose and when it has been successful in insulating employers from Title VII civil rights complaints.

The variables we coded for the case analysis are shown in table 7. The dependent variable of interest is whether a court deferred to an organization's internal grievance procedure. We coded deferral in two different ways. We coded a case as involving a "theoretical deferral" where, in the abstract (i.e., not considering the merits of the grievance procedure in that particular case), the court stated that it would base its decision at least in part on whether the organization had an internal grievance procedure.[26]

---

[24] When a court renders an opinion, it may order that the opinion be published in official case reporters; it may make the opinion generally available for public distribution but with the limitation that it is to be considered legally "unpublished" and thus not cited as precedent; or it may simply file the opinion with the general public records available at the court, in which case it generally does not appear in the official reporters or on-line databases. (On rare occasions an opinion may be "sealed" and not available to the public at all.) The Westlaw federal database is based on extensive efforts to include all "published" and (legally) "unpublished" cases from federal courts. Since in some cases a court allows the (legally) unpublished cases to appear on-line and not in the reporters, the Westlaw database is more comprehensive than the reporters. In addition, although the Westlaw database does not generally include cases where the court simply files the opinion, occasionally even those cases will be included if a request is made and the court approves it.

[25] To the extent that published cases are not representative of all cases, they tend to overrepresent novel decisions and underrepresent established law. This might mean that cases deferring to grievance procedures are somewhat overrepresented. On the other hand, all cases involving hostile work environment claims in general, and grievance procedures in particular, address novel issues and are therefore likely to be included.

[26] The coding instruction for theoretical deferral was: "In the abstract, the court would be willing to be influenced by whether the organization has an internal grievance procedure. This does not mean that the court would base its decision entirely on the presence of internal grievance; rather it simply means that the internal grievance has some degree of influence. Cases where the court says that the grievance procedure in this case is inadequate (perhaps due to lack of notice or the complainant having to complain to her supervisor) may still be coded as theoretical defers if the court would have deferred without those deficiencies in the grievance procedure."

TABLE 7

CASE ANALYSIS: VARIABLES, DESCRIPTIVE STATISTICS, AND VARIABLE DEFINITIONS

| | Mean | SD | Description |
|---|---|---|---|
| Theoretical defer ........... | 87 | .340 | "1" if court would consider deferring to an organization's grievance procedure, "0" if other |
| Actual defer . .................. | .36 | .481 | "1" if court deferred to an organization's grievance procedure, "0" if other |
| Theoretical quality ..... .. | .81 | .400 | "1" if the court held that the quality of an organization's grievance procedure ought to affect the court's decision to defer, "0" if other |
| Actual quality .. .. .. ..... | .60 | .491 | "1" if the court's decision with respect to deferral depended in part on the quality of the organization's grievance procedure, "0" if other |
| Direct liability .... ......... | .481 | .502 | "1" if case alleges either race harassment, or hostile work environment–only sex harassment, "0" if other |
| Meritor cited ................. | 29 | .458 | "1" if Meritor cited as precedent in the opinion, "0" if other |
| Year .. .. .. ...................... | 92.2 | 5.151 | Year of court decision |

NOTE.—$N = 129$.

We coded a case as involving an "actual deferral" where the court ruled in favor of the employer, in part based on the presence of an internal grievance procedure.[27] All cases involving actual deferrals must also involve theoretical deferrals, but the reverse is not true. The Meritor case, for example, involves a theoretical deferral but not an actual deferral since the court said that the grievance procedure in that case was deficient but that a better grievance procedure might have insulated the employer from liability (*Meritor Savings Bank v. Vinson*, 72–73).

The independent variables we use in the analyses include the year of the decision, whether the case was in a class where the direct liability standard is generally used (i.e., racial harassment and hostile work environment sexual harassment cases), whether the Meritor decision was ex-

[27] The coding instruction for actual deferral was: The court ruled in favor of the employer, in part based on the presence of a grievance procedure. This does not mean that the court based its decision entirely on the presence of internal grievance; rather it simply means that the internal grievance had some degree of influence. It includes cases where the court was influenced by the complainant's failure to use a grievance procedure.

438

TABLE 8

CASES DISCUSSING THE GRIEVANCE PROCEDURE DEFENSE

| Period | Total (%) | | 1974–79 (%) | | 1980–86 (before Meritor)* (%) | | 1986–97 (after Meritor)* (%) | |
|---|---|---|---|---|---|---|---|---|
| No theoretical defer ......... | 17 | (13) | 3 | (75) | 4 | (44) | 10 | (9) |
| Theoretical defer ... ......... | 112 | (87) | 1 | (25) | 5 | (56) | 106 | (91) |
| No actual defer .................... | 83 | (64) | 4 | (100) | 5 | (56) | 74 | (64) |
| Actual defer . ... ... . ......... | 46 | (36) | 0 | (0) | 4 | (44) | 42 | (36) |
| Total cases ............. . ..... ... | 129 | | 4 | | 9 | | 116 | |

* Three pre-Meritor 1986 cases are included in the second period. Meritor and one post-Meritor 1986 case are included in the third period

plicitly cited as precedent, and two variables measuring the court's stance toward the quality of the grievance procedure. The "theoretical quality" variable measures whether the court mentions in its decision that the grievance procedure should provide employees with an adequate and fair means of complaining. The "actual quality" variable means that the court based its decision in part on the quality (or lack thereof) of the grievance procedure in that particular case.

The 129 cases in which employers discussed the grievance procedure defense ranged over the period 1974–97. Our analysis indicates that *Meritor Savings Bank v. Vinson* has legitimated the grievance procedure defense. Since that case, many more employers raise the grievance procedure defense, and courts are far more predisposed to listen. As shown in table 8, the Westlaw federal database includes only 13 employment cases prior to Meritor in which the grievance procedure defense was even raised. Of those cases, there were six theoretical deferrals, four of which were also actual deferrals. All of the actual deferrals prior to the Meritor case were decided between 1980 and 1986. Although it is possible that these six cases were responsible for the claims of the personnel profession during the 1980s, it is unlikely since all were district court cases, which receive little publicity.[28] None of the six cases were cited in the personnel literature. Further, two of the six were not decided until 1986, although prior to Meritor.

Table 8 shows that, in contrast to the very small number of cases (13)

[28] One of those cases was the district court case that eventually led to the Supreme Court decision in Meritor (*Vinson v. Taylor,* 1980 Westlaw 100). The only case prior to 1980 was *Munford v. James T. Barnes,* 441 F.Supp. 459, which was decided in 1977.

addressing the grievance procedure defense prior to Meritor, there are 116 cases after Meritor that raise the grievance procedure defense. Nearly all of the post-Meritor cases (106, or 91%) involve theoretical deferrals.[29] Of those 106 cases, 42 (36%) involved actual deferrals. The high percentage of post-Meritor theoretical deferrals is particularly striking. The high percentage of post-Meritor deferrals, in part, reflects the increase in the number of hostile work environment sexual harassment cases, in which courts are the most likely to use the direct liability standard and therefore recognize the possibility of deferring to internal grievance procedures. Significantly, the rise in the number of hostile environment cases and judicial deferrals is itself endogenous. As plaintiffs and their lawyers bring actions in evolving areas of the law, employers and their lawyers frame their arguments so that courts will be likely to defer to their internal grievance procedures. These numbers strongly suggest that Meritor has encouraged employers and their lawyers to raise the grievance procedure defense.

Figure 2 complements table 8 by showing continuous-time trends for cases involving actual and theoretical deferrals in comparison to cases not involving a deferral. The figure shows that the number of cases addressing the grievance procedure defense is quite low until 1986 and then begins to rise. Theoretical deferrals begin to take off in the late 1980s, after the Meritor decision, and actual deferrals begin rising about four years later, reflecting the gradual institutionalization of the grievance procedure defense.

We next conducted a logistic regression analysis to examine how the year of the case, the type of case, judicial attention to grievance procedures, and the Meritor precedent combined to affect the likelihood of deferral. Because nearly all the cases after Meritor involve theoretical deferrals, and because the actual deferral variable provides a better measure

[29] The following case provides an example of theoretical and actual judicial deferral to organizational grievance procedures. In a 1992 federal district court case, *Giordano v. Paterson College of New Jersey* (804 F.Supp. 637 [1992]), a female campus police officer had filed both formal and informal complaints of sexual harassment naming three male coworkers. The affirmative action officer found "inappropriate behavior" but no sexual harassment. The officer recommended that all department employees attend a sexual harassment workshop, but the lawsuit alleges that the harassment continued, eventually causing the employee to quit her job. She filed suit for sexual harassment based on a hostile environment. The court held that "the College promptly and adequately responded to Giordano's complaints. As a result the College cannot be liable for sexual harassment based on a hostile environment claim" (*Giordano v. Paterson College of New Jersey*, 644). Thus, the court granted the employer's motion for summary judgement, essentially deciding that because the employer had a grievance procedure, the employee filed a grievance, and the employer took some action, the employer won as a matter of law—the case should not go to a jury for determination of whether the response of the college was adequate.
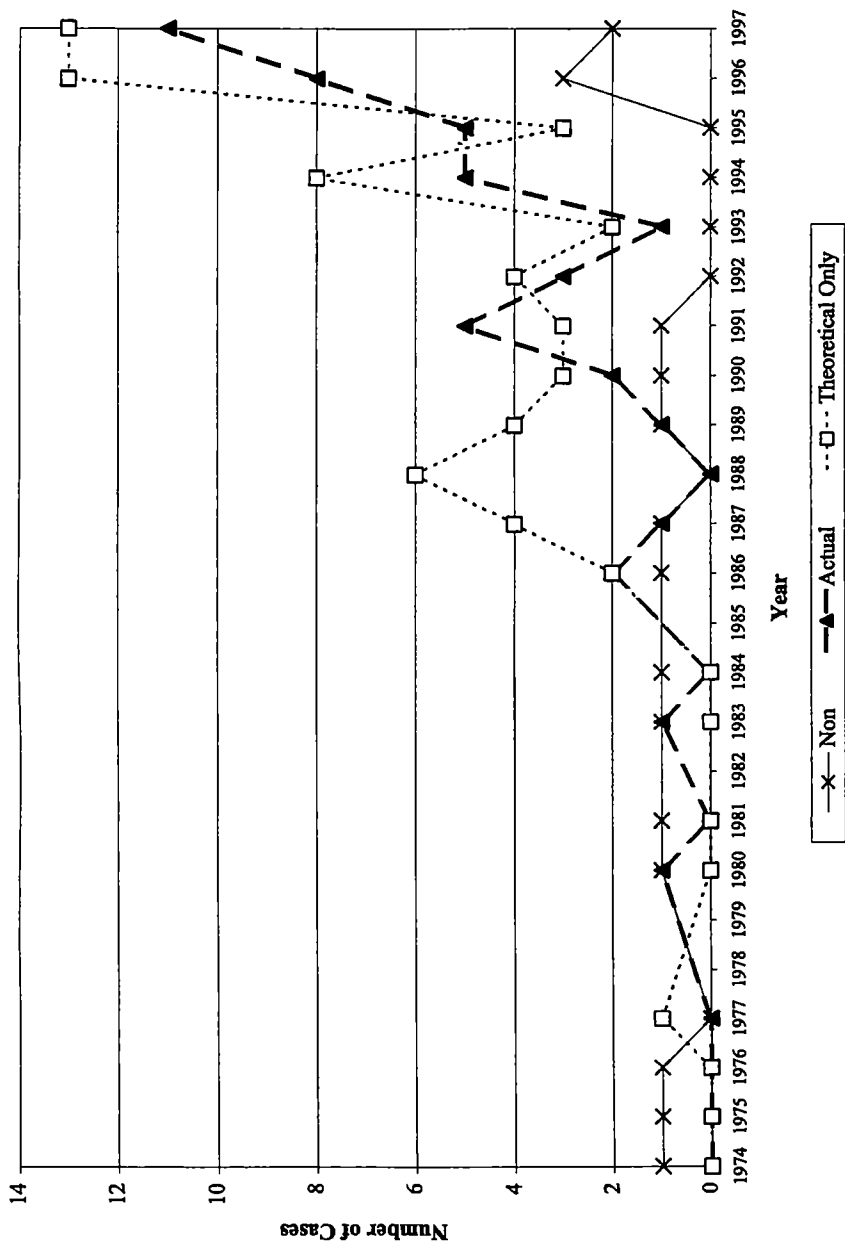
FIG. 2.—Nondeferrals, actual deferrals, and theoretical-only deferrals by year

of the success of the grievance procedure defense, we use actual deferral as the dependent variable. And we use actual quality rather than theoretical quality as an independent variable since that variable better captures judicial attention to the quality of an organization's grievance procedure. Table 9 shows logit models predicting the log-likelihood of courts' actual deferrals to internal grievance procedures. Models 1 and 2 show the likelihood of judicial deferral for all cases in the sample; these models use only the year and actual quality variables. Models 3 and 4 show the likelihood of judicial deferral for post-Meritor cases only; for these models, we add the direct liability and Meritor cited variables, which are applicable only after Meritor.

Model 1, the basic model, shows a statistically significant positive effect of time on the likelihood of actual deferrals, which suggests that courts are becoming increasingly likely to defer to internal organizational grievance procedures in making decisions about liability. This is the critical model supporting our argument about endogeneity: courts are accepting—and reinforcing—stories about the rationality and legality of organizational grievance procedures.

Models 2–4 provide additional information about this time trend by identifying the types of cases in which deferrals are most likely to occur. Model 2, which still applies to the entire sample, shows that in making these decisions courts are paying (at least some) attention to the quality of internal grievance procedures. Although courts are unlikely to have knowledge about the key ways in which internal organizational grievance procedures differ from legal forums (Edelman et al. 1993), this finding suggests that courts at least attempt to follow Meritor's guidance that poor or sham grievance procedures cannot offer protection to employers.

Model 3 replicates model 2 for the post-Meritor cases only, and model 4 adds direct liability and whether Meritor is cited to the model. In the post-Meritor years, the year variable is statistically significant even with the control variables, suggesting a continuing institutionalization of courts' tendency to defer to internal organizational grievance procedures. Judicial attention to the quality of the internal grievance procedure also significantly predicts actual deferrals.[30]

Model 4 adds the direct liability variable, which measures whether the case was a hostile work environment or racial harassment case, in which

[30] Cross-tabulations reveal that of the 78 cases where courts paid attention to the quality of the grievance procedure, there were 41 actual deferrals, whereas out of the 51 cases in which quality was not considered, there were only five actual deferrals. Moreover, all of those 41 deferrals were in cases where courts found the grievance procedures to be good. Since the courts' views of the quality of the grievance procedure were perfectly correlated with deferral, we could not use a variable measuring whether actual quality was considered and found to be good in the logit analyses.

## TABLE 9

### LIKELIHOOD OF ACTUAL JUDICIAL DEFERRAL TO ORGANIZATIONAL GRIEVANCE PROCEDURES

| VARIABLE | ALL CASES | | | | POST-MERITOR CASES | | | |
|---|---|---|---|---|---|---|---|---|
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | |
| | Coefficient | t-ratio | Coefficient | t-ratio | Coefficient | t-ratio | Coefficient | t-ratio |
| Constant | −7.767 | −1.989** | −8.393 | −1.803* | −19.063 | −2.822*** | −19.667 | −2.733*** |
| Year | .076 | 1.845* | .067 | 1.346 | .175 | 2.471** | .179 | 2.399** |
| Actual quality | | | 2.273 | 4.323*** | 2.835 | 4.283*** | 2.790 | 4.137*** |
| Direct liability (vs. vicarious liability) | | | | | | | .791 | 1.668* |
| Meritor cited | | | | | | | −.531 | −1.511 |
| Log-likelihood | −82.119 | | −69.343 | | −59.320 | | −57.148 | |
| N | 129 | | 129 | | 116 | | 116 | |

* P < .10
** P < .05
*** P < .01

the direct liability standard applies (as opposed to a discrimination case or a quid pro quo sexual harassment case in which the vicarious liability standard applies). From a legal standpoint, only the direct liability standard gives courts the option of deferring to organizational grievance procedures since the vicarious liability standard holds that employers are liable for the acts of their agents irrespective of their knowledge of those acts. The variable attains only borderline statistical significance, suggesting that while deferrals are more likely in cases using the direct liability standard, this variable does not completely explain the time trend. The result must be interpreted with caution, however, because it was not clear whether the direct liability standard was used in sexual harassment cases involving mixed hostile work environment and quid pro quo claims.[31]

The Meritor-cited variable is negative, as expected, but not statistically significant. Thus, courts that cite Meritor are not significantly less likely to defer to organizational grievance procedures when attention to the quality of the grievance procedure is statistically controlled. This makes sense, since courts are likely to cite Meritor when they pay attention to the quality of the grievance procedure, whether or not they find it of sufficiently good quality to support a deferral.

Overall, tables 8 and 9 support our contention that judicial deferral to organizational grievance procedures takes place primarily in the 1990s, quite a few years after the personnel profession's initial claims of the value of internal grievance procedures. Further, although courts' recent tendencies to defer to internal grievance procedures can be explained in part by the introduction of direct liability principles into employment law, this does *not* mean that the law alone explains the new "truth" to the personnel professions' claims about the value of internal grievance procedures. There is nothing about direct liability principles that *requires* courts to defer to internal grievance procedures. Rather, those principles simply say that an employer who knew, or should have known, or approved of the acts of an agent may be held responsible. While the conceptual leap from direct liability theory to the relevance of organizational grievance procedures is not a huge one, it was greatly facilitated by the institutionalization of nonunion grievance procedures in organizations. That institutionalization, itself a response to the personnel professions' assumptions and claims about law, made grievance procedures seem like the most rational—and indeed the most natural mode—by which employers ought to find out about problems within their organizations.

[31] There were actual deferrals in only two cases that were quid pro quo–only cases and thus clearly used the vicarious liability standard. In cases involving both quid pro quo and hostile environment claims, it was often not clear whether the court used the direct or vicarious liability standard; we coded these cases "0" on the direct liability dummy variable.

The flexibility and ambiguity inherent in the agency principle is evident in a passage from the Meritor opinion, in which the court relies on an amicus curiae brief of the EEOC, stating:

> If the employer has an expressed policy against sexual harassment and has implemented a procedure specifically designed to resolve sexual harassment claims, and if the victim does not take advantage of that procedure, the employer should be shielded from liability absent actual knowledge of the sexually hostile environment (obtained, e g., by the filing of a charge with the EEOC or a comparable state agency). In all other cases, the employer will be liable if it has actual knowledge of the harassment or if, considering all the facts of the case, the victim in question had no reasonably available avenue for making his or her complaint known to appropriate management officials. (*Meritor Savings Bank v. Vinson,* 26; citing brief for United States and EEOC as amici curiae)

That the EEOC must make a case for how agency principles should be interpreted in the context of hostile environment cases—and that Meritor cites the EEOC brief in its opinion on what agency principles mean—show that there agency principles do not *require* courts to defer to internal grievance procedures. Instead, employers like Meritor Savings Bank raise the grievance procedure defense because of the institutionalized belief that grievance procedures constitute fair treatment and that they may insulate employers from liability. Both the EEOC and the Meritor court move from the legal language involving agency principles to the importance of organizational policies and grievance procedures because of the same institutionalized beliefs. In this way, organizational policies and practices influence the construction of law in organizational settings. Courts give weight to organizational grievance procedures because ideologies of rationality suggest that these procedures constitute a reasonable alternative to litigation.

## RATIONAL MYTHS AND THE ENDOGENEITY OF LEGAL REGULATION

To unravel the endogeneity of legal regulation, we have examined the rationalization of grievance procedures with data from the professional personnel literature, the courts, and a national sample of organizations. Over the past 30 years, EEO grievance procedures have become a standard and rationalized form of compliance with EEO law. The grievance procedure example illustrates the argument that ambiguous law generates strategic attempts at rational compliance, which cannot be viewed apart from institutionally constructed belief systems. The professions, influenced by the public legal order, construct stories about what forms of compliance will minimize interference by legal institutions. Strategically,

rational organizations adopt these structures but often inject managerial goals into the legal form to make it better fit their objectives (Edelman et al. 1993).

Our analyses of the impact of organizational grievance procedures on legal complaints, and of judicial trends in deferring to organizational grievance procedures, show that the claims made by the personnel profession regarding the insulating power of grievance procedures were largely unfounded at the time they were made. Organizational grievance procedures appear to have no impact on the number of complaints to legal agencies and in fact increase the number of internal complaints. They may result in diversion of complaints from external to internal forums but probably only because EEO offices present in the same organization encourage employees to file more complaints and different types of complaints internally.

In addition, organizational grievance procedures had little legal value until recently and then primarily in cases where the direct liability standard has been adopted. Thus, the stories that the personnel profession created and disseminated about the rationality of organizational grievance procedures were not unreasonable ones; it took only small leaps in legal logic to believe that grievance procedures could provide insulation from the legal environment. But in fact, there was almost no basis in the law for these claims. It seems likely that the personnel professionals who were publishing on the topic simply "knew" the value of grievance procedures—because grievance procedures have been legitimated in organizational and socio-legal fields as institutions of legality and justice. Further, because the professions act as carriers of important information about the legal environment (Edelman 1992), organizations adopted grievance procedures without serious review of these claims.

The fictive status of these rational myths has now begun to change, however, because courts—as part of the same organizational and socio-legal fields—are influenced by the same set of institutionalized beliefs as other organizations (Edelman, in press). And as courts begin to legitimate grievance procedures, the professions also step up their message.[32] Even since Meritor, the professions appear to construct straightforward stories about the rationality of grievance procedures, rather than delving into the

[32] Advice on the value of grievance procedures is not confined, moreover, to professional personnel journals or books; it is easily found on numerous sites on the World Wide Web. The Workforce Online site, for example, has a page titled "Minimize Risk by Investigating Complaints Promptly" The site advises employers that "Recent wrongful discharge and sexual harassment verdicts underscore an employer's liability for failing to provide for an effective means to resolve sexual and other harassment issues in the workplace. . . . To avoid such risks, employers must develop and implement an effective method for employees to raise sexual harassment complaints."

complexities and contradictions in the court decisions. For example, Alfred Feliu, shortly after the Meritor decision, wrote, "The existence of an internal system enhances both the appearance and reality of fairness of the company's actions in a later court proceeding. . . . The [Meritor] Court viewed the existence of an effective internal complaint procedure as a positive—and possibly dispositive—factor in defense of sexual harassment (and presumably other discrimination) claims" (Feliu 1987, pp. 91–92).

The Feliu quote, which appeared just after Meritor was decided, overstates the legal value of grievance procedures by suggesting that grievance procedures may be a "dispositive" factor in sexual harassment cases and "presumably in other discrimination cases," when in fact it is almost never dispositive and (even today) would probably not be an important factor in other discrimination cases.

The grievance procedure example, then, highlights the endogeneity of court decisions, professional norms, and organizational practices. Organizations create grievance procedures as evidence of fair treatment, and the professions make claims about the legal value of those procedures. As those claims become institutionalized, forming ideologies of rationality, courts recognize and construct the rationality of EEO grievance procedures by holding that, under some circumstances, these procedures constitute evidence of fair treatment and that employers may escape liability by adhering to their (effective) grievance resolution procedures. The professions, sometimes with greater enthusiasm than is perhaps warranted, filter and disseminate court decisions, which reinforce and legitimate organizations' initial structural responses to law. And the circle closes as organizations continue and elaborate their responses.

Of course, the organizational and legal treatment of grievance procedures has developed in the context of a highly ambiguous, politically contested, and weakly enforced set of laws. Had the law been clearer about the meaning of discrimination, or less politically contested, courts might defer less to the procedures and decisions of organizational actors. And, it may be the case that where expected enforcement is very weak, organizations may not even find it important to make symbolic gestures toward—or attempt to insulate themselves from—the legal environment. In fact, the pattern of EEO grievance procedure adoptions we find would be consistent with this argument since very few procedures appear until the early 1970s, a time at which EEO enforcement became somewhat more stringent.

Although we have focused on the development of the grievance procedure defense in Title VII actions, it is important to recognize that judicial deferral to organizational dispute resolution forums is not limited to federal discrimination cases. A 1996 case involving wrongful discharge in violation of an implied contract decided in the California court of appeals (*Cotran v. Rollins Hudig Hall International, Inc.,* 57 Cal.Rptr.2d 129

[1996]), for example, shows that judicial deferral to organizational forums is also occurring in implied contract wrongful discharge cases in state courts.[33] The Cotran case is especially important because the court says explicitly that the result of an internal hearing need not be correct as long as the employer is acting in good faith. "To require an employer to be 'right' about the facts on which it bases its decision to terminate an employer is to interfere with the wide latitude an employer must have in making independent, good faith judgements about high-ranking employees without the threat of a jury second-guessing its business judgement. . . . Where, as here, we are dealing with charges of sexual harassment, we believe the most we can reasonably ask of employers under these difficult circumstances is that they act responsibly and in good faith" (*Cotran v. Rollins Hudig Hall International, Inc.,* 139–40).

Courts also defer, under some circumstances, to mandatory arbitration clauses. In 1991, the United States Supreme Court upheld *arbitration* clauses under the Federal Arbitration Act in *Gilmer v. Interstate/Johnson Lane Corp.* (111 Sup. Ct. 1647 [1991]). The court held that arbitration was no longer just a private means of resolving disputes between freely contracting parties but also an alternative forum for resolution of public statutory rights. Since then, a majority of the federal appellate circuits have held that Gilmer extends to mandatory arbitration of discrimination claims under employment contracts, and many employers are inserting mandatory arbitration clauses into employment contracts (Matthews 1997). And as in the case of organizational grievance procedures, numerous personnel articles and Web sites tout the rationality of mandatory arbitration.

Legislatures are also in some sense deferring to organizational realms. Both the 1990 Americans with Disabilities Act and the 1991 Civil Rights Act encourage employers to establish or participate in forms of alternative dispute resolution as a means of resolving complaints without resorting to litigation.

From a policy standpoint, judicial deferral to organizations—and the endogeneity of law generally—raises some concerns. Although Meritor warned that courts should defer only to effective grievance procedures, and our data suggest that many courts are heeding that warning, courts

---

[33] Ralph Cotran was fired from Rollins Hudig Hall for sexual harassment after two women used an internal grievance procedure to file sexual harassment charges against him and an internal investigation held that it was more likely than not that sexual harassment had occurred. Cotran then filed suit for wrongful termination claiming that there was an implied contract that he be terminated only for good cause. The jury in his wrongful termination suit found Cotran had not engaged in sexual harassment and awarded him $1,783,549 for wrongful termination. The California court of appeals, however, overturned the verdict, finding that the lower court should not have overturned the findings of the organization's internal grievance procedure.

are not likely to be aware of the many ways in which internal grievance forums may undermine legal rights. Even when employers have the best of intentions, the lack of formal due process protections that characterize internal grievance procedures, the different types of remedies offered by those forums, and the tendency of organizational complaint handlers to recast legal problems in managerial terms mean that legal rights are often transformed in the organizational context (Edelman et al. 1993). Further, employees may have good reasons for choosing not to use internal grievance procedures; for example, they may be legitimately concerned about biases of decision makers or retaliatory tactics of employers but may be unable to demonstrate these problems to a court. To the extent that courts defer to organizations without full awareness of the subtle pressures that characterize the workplace, legal ideals may be compromised. More generally, when courts adopt forms of compliance created within organizational fields, they run the risk of institutionalizing the very forms of discrimination that laws were originally designed to alleviate.

From a theoretical standpoint, our analysis of grievance procedures suggests an important corrective to the organizational literature on rationality. Although much of the literature debates whether rationality is "real," in the sense of providing market benefits (e.g., Williamson 1975), or socially constructed (e.g., Meyer and Rowan 1977), our study suggests that socially constructed rationality may over time produce market benefits as courts recognize and legitimate organizational practices and hold that they may protect organizations from liability.

Our example also elaborates the socio-legal literature on the symbolic nature of law (Edelman 1992; Stryker 1990, 1996) by illustrating how legal rules function less as absolute mandates than as rhetorical claims that acquire meaning through dialog across organizational, professional, and legal fields. That dialog operates as ideologies of rationality evolve within and flow across those fields, influencing ideas of justice as well as efficiency. Developments within each field spur reactions, interpretations, normative declarations, filtering processes, and organizational mimesis, all of which interact to construct not only organizational responses to law but the law itself.

APPENDIX

The Diffusion of EEO Grievance Procedures

Event-history models of the diffusion of EEO grievance procedures over time suggest that organizations did in fact heed the warnings and advice of the personnel profession. We estimated models of the form:

$$r_p = \alpha_p e^{\beta x + \lambda_p z_p},$$

TABLE A1

DETERMINANTS OF THE CREATION OF SPECIFIC GRIEVANCE PROCEDURES

| Variables | Coefficient | Antilog | $t$-value |
|---|---|---|---|
| Time-independent variables: | | | |
| Sector: | | | |
| College ................................ . . | .75*** | 2.12 | 2.75 |
| Government (business) ........... | .94** | 2.56 | 2.17 |
| Log size ..... . ... ........................ | 1.47* | 1.15 | 1.89 |
| 1964–76: | | | |
| Constant .................................... .... | −6.63*** | .001 | −11.81 |
| Contractor ..... .. ........................... | .91** | 2.49 | 2.05 |
| EEO office ................................... | 1.24** | 3.45 | 2.52 |
| Lawsuit ............................. .. ... ... . | −8.29 | .0002 | −.18 |
| 1977–79: | | | |
| Constant ..................................... | −4.92*** | .007 | −9.31 |
| Contractor ......... .. ...... .. ... .. ....... | .29 | 1.34 | .51 |
| EEO office .............................. .. .. | 1.18** | 3.24 | 2.36 |
| Lawsuit ..................................... ... | −1.08 | .34 | −1.36 |
| 1980–89: | | | |
| Constant ...... ... .. . .. .. ................. | −5.49*** | .004 | −11 61 |
| Contractor ............................. ..... | −.39 | .68 | −1.09 |
| EEO office ................................... | .56 | 1.75 | 1 36 |
| Lawsuit . ..................................... | .93** | 2.54 | 2 54 |

* $P < .10$.
** $P < .05$.
*** $P < .01$.

where $r_p$ represents the formation rate from the starting state (no structure) to the ending state (the creation of an internal EEO procedure) during period $p$; $\alpha_p$ represents the constant term for period $p$; $\beta x$ represents a vector of explanatory variables with effects that are not dependent on time period and their coefficients; and $\lambda_p z_p$ represents a vector of time-dependent explanatory variables and their coefficients.[34] Based on chi-square likelihood ratio tests of various period specifications, we allow the rate of procedure creation to vary across three historical periods: 1964–74, 1975–79, and 1980–89.

Table A1 in the appendix shows an event-history model of the diffusion of EEO grievance procedures. The coefficients give the effect of the vari-

---

[34] Event-history analysis is appropriate for modeling the sources of EEO grievance procedure creation because legal environments exert continuous pressure on organizations (Tuma and Hannan 1984; Allison 1984; Carroll 1983). The dependent variable is the rate of EEO grievance procedure creation, where the rate is defined as "the transitional probability over a unit of time where the unit is infinitesimal" (Carroll 1983).

able on the log of the rate of special grievance procedure formation, while the antilogs of the coefficients give the multipliers for the base rate, or the relative risk of an EEO grievance procedure. The results for EEO grievance procedures are consistent with earlier analyses of the diffusion of institutional models, which addressed nonunion grievance procedures and general grievance procedures (Edelman 1990; Sutton et al. 1994). Organizational sector exerts a relatively strong effect on the rate of EEO grievance procedure formation. Government organizations are almost twice as likely as business organizations to create EEO grievance procedures, and colleges are also somewhat more likely than businesses to create such procedures. Organizational size has a marginally statistically significant effect on EEO grievance procedure formation, which may represent the greater visibility (Edelman 1992) or the greater financial and technical resources (Yeager 1991) of larger organizations.

During the first time period, federal contractor status and the presence of an EEO office predict EEO grievance procedure formation: both render organizations far more sensitive to their legal environments, and EEO offices provide a vehicle for the infusion of professional norms. Contractor status is no longer statistically significant during the 1980s, a period of waning civil rights enforcement. During the later period, experiencing a discrimination lawsuit is a positive predictor of EEO grievance procedure. The late effect of lawsuits may suggest that lawsuits motivate organizations to adopt institutionalized symbolic structures; if lawsuits motivated internal grievance procedures purely as a cost-reducing measure, one would expect a positive effect in the earlier periods as well.

The event-history findings, then, show that factors rendering organizations more sensitive to their legal environments—proximity to the public sphere, federal contractor status, the presence of an EEO office, and previous lawsuit experience—motivate a higher rate of EEO grievance procedure formation, especially during the late 1960s and 1970s when those procedures were spreading rapidly among organizations. These findings suggest that institutional environments provide an important source of ideas regarding the rationality of grievance procedures. The admonitions and advice of the legal and personnel professions, discussed above, illustrate an important mechanism by which these environments operate.

## REFERENCES

Abbott, Andrew. 1988. *The System of Professions.* Chicago: University of Chicago Press.

Allison, Paul D. 1984. *Event History Analysis: Regression for Longitudinal Event Data.* Beverly Hills, Calif.: Sage.

Baron, James N., Frank R. Dobbin, and P. Devereaux Jennings. 1986. "War and Peace: The Evolution of Modern Personnel Administration in U.S. Industry." *American Journal of Sociology* 92:350–83.

Becker, Gary. 1971. *The Economics of Discrimination,* 2d ed. Chicago: University of Chicago Press.

Berger, Peter L., and Thomas Luckmann. 1967. *The Social Construction of Reality.* New York: Doubleday

Carroll, Glenn R. 1983. "Dynamic Analysis of Discrete Dependent Variables: A Didactic Essay." *Quality and Quantity* 17:425–60.

Caspi, Avshalom, Bradley R. Entner-Wright, Terrie E. Moffitt, and Phil A. Silva. 1998. "Childhood Predictors of Unemployment in Early Adulthood." *American Sociological Review* 63:424–51.

Condon, Thomas J. 1985. "Use Union Methods in Handling Grievances." *Personnel Journal* 64:72–75.

Corzine, James E. 1967 "Structure and Utilization of a Grievance Procedure." *Personnel Journal* 46:484–89.

DiMaggio, Paul J. 1988. "Interest and Agency in Institutional Theory." Pp. 3–21 in *Institutional Patterns and Organizations,* edited by Lynne G. Zucker. Cambridge, Mass.: Ballinger.

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48:147–60.

Dobbin, Frank R., Lauren B. Edelman, John W. Meyer, W. Richard Scott, and Ann Swidler. 1988. "The Expansion of Due Process in Organizations." Pp. 71–98 in *Institutional Patterns and Organizations: Culture and Environment,* edited by Lynn G. Zucker. Cambridge, Mass.: Ballinger.

Dobbin, Frank R., John R. Sutton, John W. Meyer, and W. Richard Scott. 1993. "Equal Employment Opportunity Law and the Construction of Internal Labor Markets." *American Journal of Sociology* 99:396–427.

Donahue, John J., III, and Peter Siegelman. 1991. "The Changing Nature of Employment Discrimination Litigation." *Stanford Law Review* 43:983–1033.

Edelman, Lauren B. 1990. "Legal Environments and Organizational Governance: The Expansion of Due Process in the Workplace." *American Journal of Sociology* 95:1401–40

———. 1992. "Legal Ambiguity and Symbolic Structures: Organizational Mediation of Law." *American Journal of Sociology* 97:1531–76.

———. In press. "Constructed Legalities: Socio-Legal Fields and the Endogeneity of Law." In *Bending the Bars of the Iron Cage: Institutional Dynamics and Processes,* edited by Walter W. Powell and Daniel Jones. Chicago: University of Chicago Press

Edelman, Lauren B., Steven E. Abraham, and Howard S. Erlanger. 1992. "Professional Construction of the Legal Environment: The Inflated Threat of Wrongful Discharge Doctrine." *Law and Society Review* 26:47–83.

Edelman, Lauren B., Howard S. Erlanger, and John Lande. 1993. "Employers' Handling of Discrimination Complaints: The Transformation of Rights in the Workplace." *Law and Society Review* 27:497–534.

Edelman, Lauren B., and Stephen Petterson. 1999. "Symbols and Substance in Organizational Response to Civil Rights Law." In *Research in Social Stratification and Mobility,* vol. 17. Special issue on the future of affirmative action. In press.

Edelman, Lauren B., Stephen Petterson, Elizabeth Chambliss, and Howard S. Erlanger. 1991. "Legal Ambiguity and the Politics of Compliance: Affirmative Action Officers' Dilemma." *Law and Policy* 13:73–97.

Epstein, Richard A. 1992. *Forbidden Grounds: The Case against Employment Discrimination Laws.* Cambridge, Mass.: Harvard University Press.

Ewing, David W. 1982. "Due Process: Will Business Default?" *Harvard Business Review* 60:114–22.

Feliu, Alfred G. 1987. "Legal Consequences of Nonunion Dispute-Resolution Systems." *Employee Relations Law Journal* 13:83–103

Fuller, Sally Riggs. 1993. *Symbolic Actions and Antics: Mechanisms for Shaping Organizational Reality*. Doctoral dissertation. University of Wisconsin—Madison, School of Business.

Galanter, Marc. 1983. "Reading the Landscape of Disputes: What We Know and Don't Know (and Think We Know) about Our Allegedly Contentious and Litigious Society." *UCLA Law Review* 31:4–71.

———. 1994. "Predators and Parasites: Lawyer-Bashing and Civil Justice." *Georgia Law Review* 28:633–81.

Gibbs, Jack P., and Glenn Firebaugh. 1985. "User's Guide to Ratio Variables." *American Sociological Review* 50:713–22.

Heckman, James J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5:475–92.

Heckman, James J., and Kenneth I. Wolpin. 1976. "Does the Contract Compliance Program Work? An Analysis of Chicago Data." *Industrial and Labor Relations Review* 29:544–64.

Hipp, Charles E., Jr. 1988. "Now You See It, Now You Don't: The 'Hostile Work Environment' after Meritor." *American Business Law Journal* 26:339–61.

Jacoby, Sanford M. 1985. *Employing Bureaucracy: Managers, Unions, and the Transformation of Work in American Industry, 1900–1945*. New York: Columbia University Press.

Konrad, Alison M., and Frank Linnehan. 1995. "Formalized HRM Structures: Coordinating Equal Employment Opportunity or Concealing Organizational Practices?" *Academy of Management Journal* 38 (3): 787–820.

Larson, Margali Sarfatti. 1977. *The Rise of Professionalism*. Berkeley: University of California Press.

Leonard, Jonathan S. 1984. "The Impact of Affirmative Action on Employment." *Journal of Labor Economics* 2:439–63.

———. 1986. "The Effectiveness of Equal Employment Law and Affirmative Action Regulation." *Research in Labor Economics* 8:319–50.

Linenberger, Patricia, and Timothy J. Keaveny. 1981. "Sexual Harassment: The Employer's Legal Obligations." *Personnel* 58 (6): 60–68.

Little, Roderick J., and Donald Rubin. 1987. *Statistical Analysis with Missing Data* New York: Wiley.

Lo Bosco, Maryellen. 1985. "Nonunion Grievance Procedures." *Personnel* 62:61–64.

Maddala, G. S. 1983. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge, Mass.: Cambridge University Press.

Matthews, Donna M. 1997. "Employment Law after Gilmer: Compulsory Arbitration of Statutory Antidiscrimination Rights." *Berkeley Journal of Employment and Labor Law* 18:347–87.

McDonald, John F., and Robert A. Moffitt. 1980. "The Uses of Tobit Analysis." *The Review of Economics and Statistics* 62:318–21.

Meyer, John W., and Brian Rowan. 1977. "Institutionalized Organizations. Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83:340–63.

Mezias, Stephen J. 1990. "An Institutional Model of Organizational Practice: Financial Reporting at the Financial 200." *Administrative Science Quarterly* 31:431–57.

Oliver, Christine. 1991. "Strategic Responses to Institutional Processes." *Academy of Management Review* 16:145–79.

Panken, Peter M. 1984 "What Every Company Should Have: A Formal Employee Complaint Procedure." *Management Review* 73:42–45.

Pfeffer, Jeffrey. 1981. *Power in Organizations*. Cambridge, Mass.: Ballinger.

Pfeffer, Jeffrey, and Gerald R. Salancik. 1978. *The External Control of Organizations*. New York: Harper & Row.

Posner, Richard A. 1986. *Economic Analysis of Law*, 3d ed. Boston: Little, Brown.

Powell, Walter W. 1985. "The Institutionalization of Rational Organization: Review of Organizational Environments." *Contemporary Sociology* 14 (5): 564–66.

Powell, Walter W., and Paul J. DiMaggio, eds. 1991. *The New Institutionalism in Organizational Analysis.* Chicago: University of Chicago Press.

Roncek, Dennis W. 1992. "Learning More from Tobit Coefficients. Extending a Comparative Analysis." *American Sociological Review* 57:503–7.

Scott, W. Richard. 1983. "The Organization of Environments: Network, Cultural and Historical Elements." Pp. 129–53 in *Organizational Environment: Ritual and Rationality,* edited by John W. Meyer and W. Richard Scott. Beverly Hills, Calif.: Sage.

———. 1987. *Organizations: Rational, Natural, and Open Systems,* 2d ed. Englewood Cliffs, N.J.. Prentice-Hall.

Scott, W. Richard, and John W. Meyer. 1983 "The Organization of Societal Sectors." Pp. 129–54 in *Organizational Environments: Ritual and Rationality,* edited by John W. Meyer and W. Richard Scott. Beverly Hills, Calif.: Sage.

Sewell, William H., Jr. 1992. "A Theory of Structure: Duality, Agency, and Transformation." *American Journal of Sociology* 98:1–29.

Stryker, Robin. 1990. "A Tale of Two Agencies: Class, Political-Institutional, and Organizational Factors Affecting State Reliance on Social Science." *Politics and Society* 18 (1): 101–41.

———. 1994. "Rules, Resources, and Legitimacy Processes: Some Implications for Social Conflict, Order, and Change." *American Journal of Sociology* 99:847–910.

———. 1996. "Beyond Theory vs. History: Strategic Narrative and Sociological Explanation." *Sociological Methods and Research* 24:304–52.

Suchman, Mark. 1993. *On Advice of Counsel: Legal and Financial Firms as Information Intermediaries in the Development of Silicon Valley.* Doctoral dissertation. Stanford University, Department of Sociology.

Suchman, Mark C., and Lauren B. Edelman. 1996. "Legal-Rational Myths: Lessons for the New Institutionalism from the Law and Society Tradition." *Law and Social Inquiry* 21:903–41.

Sutton, John R., Frank Dobbin, John W. Meyer, and W. Richard Scott. 1994. "The Legalization of the Workplace." *American Journal of Sociology* 99:944–71.

Thompson, James D. 1967. *Organizations in Action* New York: McGraw-Hill.

Tobin, James. 1958. "Estimation of Relationships for Limited Dependent Variables." *Econometrica* 26:24–36.

Tuma, Nancy, and Michael T. Hannan. 1984. *Social Dynamics: Models and Methods.* Orlando, Fla.: Academic Press.

Williamson, Oliver E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications.* New York: Free Press.

———. 1979. "Transaction-Cost Economics: The Governance of Contractual Relations." *Journal of Law and Economics* 22:233–61.

———. 1981. "The Economics of Organization: The Transactions Cost Approach." *American Journal of Sociology* 87·548–77.

Winship, Christopher, and Robert D. Mare. 1992. "Models for Sample Selection Bias." *Annual Review of Sociology* 18:327–50.

Yeager, Peter. 1991. *The Limits of Law: The Public Regulation of Private Pollution.* Cambridge: Cambridge University Press.

Zucker, Lynn. 1977. "The Role of Institutionalization in Cultural Persistence." *American Sociological Review* 42:726–43.

———. 1988. "Where Do Institutional Patterns Come From? Organizations as Actors in Social Systems." Pp. 23–49 in *Institutional Patterns and Organizations: Culture and Environment,* edited by Lynn G Zucker. Cambridge, Mass.: Ballinger.

# Civil Rights Law at Work: Sex Discrimination and the Rise of Maternity Leave Policies[1]

Erin Kelly and Frank Dobbin
*Princeton University*

By the time Congress passed the Family and Medical Leave Act of 1993, many employers had created maternity leave programs. Analysts argue that they did so in response to the feminization of the workforce. This study charts the spread of maternity leave policies between 1955 and 1985 in a sample of 279 organizations. Sex discrimination law played a key role in the rise of maternity leave policies. Building on neoinstitutional theory, this article explores how the separation of powers shapes employer response to law. Details of the law are often specified in administrative rulings—the weakest link in the law because they can be overturned by the courts and by Congress. Yet an administrative ruling requiring employers with disability leave programs to permit maternity leave, which employers successfully fought in the courts, was at least as effective as the identical congressional statute that replaced it. In the American context, the legal vulnerability of administrative rulings can draw attention to them, thus making the weakest link in the law surprisingly powerful.

## INTRODUCTION

In the first debate of the 1996 presidential campaign, Bob Dole (*Washington Post* 1996, p. A12) argued that Congress should not have used the "long arm of the federal government" in 1993 to force employers to offer

parental leaves. Most employers were already providing maternity leave in response to the growing ranks of women workers, Dole argued. Business leaders had made the same point in fighting the Family and Medical Leave Act (FMLA) of 1993. James A. Klein, the manager of employee benefits for the United States Chamber of Commerce, argued: "There is a growing sense, as the demographics of our work force change, that parental leave is a good benefit to have. . . . A recent Bureau of National Affairs survey showed that 90% of companies grant maternity leave. . . . We think that's terrific, so long as it's voluntary" (quoted in Pear 1985, p. 16). In a letter to the *Washington Post*, the Small Business Association stated: "In the absence of any government mandate . . . between 74 and 90 percent of all businesses are already addressing the problem" (Jasinoski 1991, p. 3C).

It is true that employer treatment of maternity had changed before passage of the FMLA in 1993. In the late 1950s, only 5% of those in our sample of medium and large employers offered maternity leave; by 1985, over half offered it. In the late 1960s, many corporations still required pregnant women to resign; by the late 1970s, few did so (Fryburger 1975). Many women took advantage of these changes in employer policy. In the early 1960s, only one-sixth of employed women who bore children took leaves; by the mid-1980s fully half were taking leaves (O'Connell 1990).

Employers created maternity leave programs in the 1970s and 1980s, but not as a "voluntary" response in "the absence of any government mandate," as politicians, business leaders, and scholars have argued. Employers responded to three kinds of government mandates. First, in 1972, the Equal Employment Opportunity Commission (EEOC) ruled that employers who allowed leaves for disabling medical conditions must allow them for maternity. Failure to do so constituted sex discrimination under the Civil Rights Act of 1964. As early as 1973, a survey (Bender 1973, p. 61) found that 58% of large employers had responded with new maternity leave policies. Second, in 1978, after the Supreme Court overturned the EEOC ruling, Congress passed a statute codifying the 1972 EEOC requirements. Third, some states went a step further. California required employers to allow job-guaranteed maternity leaves, regardless of their temporary disability policies.

We show that these legal changes had a profound, if little-recognized, influence on workplace maternity leave (Meyer and Rowan 1977; Dobbin and Sutton 1998). Most students of maternity leave programs neglect these policies, in some cases because they use cross-sectional data that make it hard to observe the effects of legal changes (Glass and Fujimoto 1995; Goodstein 1994; Ingram and Simons 1995; Milliken, Martins, and Morgan 1998). But even those who focus on the legal environment tend to overlook early maternity leave law. Ruhm and Teague (1997, p. 133) argue, "Prior

to 1993 there was no federal law requiring U.S. employers to offer parental leave," despite the fact that 1972 and 1978 laws required many employers to offer maternity disability leaves and pay (see also Klerman and Liebowitz 1997). Like Bob Dole and the Chamber of Commerce, most scholars discount the role of public policy and privilege the role of the labor market when explaining employers' leave policies.

We build on neoinstitutional studies that explore why Americans see business practices that are driven by public policy as driven by market forces (Fligstein 1990, 1996; Dobbin and Sutton 1998). Americans tend to underestimate the importance of policy in part because the federal government appears to be weak. We argue that one basis of the federal government's perceived weakness, the separation of powers, can produce a peculiar sort of policy efficacy. In separating governing powers, the Constitution is thought to weaken the state by making administrative and case law and, to a lesser extent, legislation susceptible to legal challenges. History has created a natural laboratory that allows us to examine the relative efficacy of an administrative ruling and a parallel, subsequent, congressional statute. Our findings challenge the conventional wisdom that administrative regulations are ineffective when their legal standing is uncertain. Because the 1972 EEOC ruling was contested in court, and ultimately because the administrative branch's authority to make law is tenuous, the ruling won press attention that popularized maternity leave policies. Yet perhaps because the ruling was struck down by the Supreme Court, politicians and scholars have failed to recognize its effects.

We also provide what we believe to be the first test of the relative efficacy of equal opportunity and affirmative action laws in changing employer practices, and here our findings confirm the conventional wisdom that affirmative action laws are more effective. We test these hypotheses about the effects of public policy in the context of competing hypotheses about female labor force participation and about employer sensitivity to the risk of legal sanction. Like Bob Dole and the Chamber of Commerce, most labor economists assume that employers adopted maternity leave in response to the growing feminization of the workforce. We look at the effects of industry feminization and trends in feminization. Students of law and regulation, on the other hand, argue that employers comply with the law when the risk of legal sanction is high. We compare the effects of a contested administrative ruling of uncertain legal standing with those of an uncontested congressional statute. We also operationalize the risk of legal sanction with a measure of federal lawsuits dealing with maternity leave. The feminization and legal sanction hypotheses fare poorly in our analyses.

Our analysis relies on data from a stratified random sample of California, New Jersey, and Virginia employers. After charting the history of maternity leave law and deriving hypotheses about the effects of changes

in the law, we review theories that suggest rival hypotheses. We then model the adoption of maternity leave policies in a sample of 279 organizations, 148 of which established formal leave policies between 1955 and 1985. We look at the adoption of written maternity leave policies, which guarantee that a worker can return to her job, or to a similar job, after being away from work due to pregnancy, childbirth, or infant care.[2]

## TRANSLATING LAW INTO ORGANIZATIONAL RULES

Classical organizational theorists paid scant attention to the role of the state in shaping the behavior of organizations. Neoinstitutional theorists have remedied this, showing that a variety of contemporary organizational features had their origins in public policy. Fligstein (1990) shows that antitrust amendments encouraged firms to adopt the conglomerate form. Roy (1997) shows that the modern private corporation inherited the structure of the 19th-century, public-purpose corporation. Baron, Dobbin, and Jennings (1986) show that World War II labor market regulations encouraged firms to develop personnel systems. Before these neoinstitutional studies, the role of the state had faded in scholarly accounts of corporate attributes, arguably because antistatist sentiments led Americans to articulate market rationales for all corporate practices.

Neoinstitutionalists have sought to understand the central paradox of U.S. business regulation—how a seemingly weak state can shape corporate practice. While some theorists had described the relationship between the state and organizations as "coercive" (DiMaggio and Powell 1983), empirical studies found that the U.S. state often influences corporate behavior without using a clear, coercive mandate. Neoinstitutionalists first explored why firms respond to ambiguous legislation. They found that ambiguous federal laws could produce elaborate compliance activities among organizations uncertain of what is required of them. Ambiguity in legislation leads firms to invent compliance measures on their own, to be tested by the courts (Edelman 1992; Dobbin et al. 1993). The measures they invent feed into the next round of statute making (Carruthers and Halliday 1998, p. 45). Professional groups—lawyers and personnel managers—turned legal ambiguity into a professional asset, by exaggerating the risk of litigation to win corporate resources (Edelman et al. 1992) and by lobbying executives to create extensive, legalistic personnel systems (Sutton and Dobbin 1996).

---

[2] Over the past 40 years, maternity leave has been linked to disability leave, sick leave, parental leave, and family leave. We consider policies covering maternity to be maternity leave policies, regardless of the label used.

The Paradox of Administrative Law

Our primary theoretical goal is to explore the effects of another apparent weakness of the U.S. state, the separation of powers. The separation of powers permits all three branches of government to make law, but it also opens the laws made by each branch to challenge. Congress writes laws, often using abstract language that demands interpretation. Federal agencies then interpret the law, issuing regulations that begin to define compliance. The courts can rule against these regulations on the grounds that they are inconsistent with legislation or with constitutional law. Affected parties often challenge administrative and judicial rulings, and free riders often await final disposition of the challenges brought by their peers. Congress also has the power to write laws that directly address case law and administrative law, validating some interpretations of previous statutes and overruling others. In short, case law trumps administrative law, and legislation trumps everything but the Constitution.

Most see the separation of powers as weakening the state, because it renders federal edicts vulnerable to legal challenge and reversal. It makes administrative law particularly vulnerable. Scholars of law and regulation from both left and right (Sunstein 1996; Posner 1997) suggest that corporate responsiveness to the law is a function of the risk of sanction. The likelihood of sanction in the federal courts is greater in the case of congressional acts than in the case of administrative law, because legislation is less susceptible to legal challenge and reversal. It follows that, all else being equal, legislation should be more effective than administrative law at changing employer practices.

The history of early maternity leave law sheds light on the process by which law—both statutes and administrative rulings—are translated into organizational practices. With the Civil Rights Act of 1964, Congress outlawed sex discrimination in employment. The EEOC, established to administer the law, at first defined sex discrimination narrowly. In 1972, it expanded the definition to include pregnancy discrimination. That definition was successfully challenged in court by General Electric and was then revived by Congress in the Pregnancy Discrimination Act of 1978. We use this historical sequence to assess the relative efficacy of administrative law. We compare the contested administrative ruling with the virtually identical congressional act that replaced it. Both the 1972 ruling and the 1978 act required employers to treat pregnancy-related disability on par with other short-term disabilities—to offer leaves and pay for pregnancy if they offered them for other medical conditions. Despite the fact that it was overturned by the Supreme Court, the 1972 EEOC ruling was surprisingly effective at changing employer policies.

We argue that in the United States, where employment law is enforced

reactively through litigation rather than proactively through administrative scrutiny, employers respond to the visibility of new laws and the perceived risk of litigation rather than to the objective risk of legal sanction. Press attention to legal battles increases the perceived risk of sanction, in part because it increases the visibility of the law and in part because negative publicity is a potent sanction itself. In other words, employment law may work through such informal sanctions as bad press as well as through such formal sanctions as costly litigation (Suchman and Edelman 1996; McCann 1998). Thus litigation and press coverage can serve to promulgate a new federal edict, even when the goal of the litigation is to overturn the edict. Paradoxically, then, administrative rulings may elicit organizational response *because* they are inherently susceptible to court challenges.

Affirmative Action versus Equal Opportunity Laws

Our second goal is to examine the relative effectiveness of equal opportunity and affirmative action laws. Equal opportunity laws require employers to treat all employees equally, regardless of characteristics such as sex, race, and religion. Affirmative action law, as Kennedy and Johnson defined it, encourages positive efforts to improve opportunities for disadvantaged groups (see Skrentny 1996). Activists within the Civil Rights and women's movements have long debated the merits of affirmative action in employment. While many contend that it improves opportunities for disadvantaged groups, others contend that it smacks of "special treatment" and thus lacks the legitimacy conferred by the 14th amendment's language of equal protection.

Few previous studies have explored the efficacy of affirmative action laws in the realm of employment. Early labor market studies showed that employers covered by affirmative action law were more likely to add black men to their payrolls than were employers covered only by equal opportunity law (Ashenfelter and Heckman 1976; Leonard 1984). Like the authors of these early labor market studies, we compare two groups of employers: one exposed to an equal treatment provision and the other exposed to both that provision and an affirmative action law. The federal Pregnancy Discrimination Act of 1978 required equal treatment of pregnancy and other disabling medical conditions. California's 1978 "affirmative action" maternity law additionally required employers to offer job-guaranteed maternity leave. We find that the affirmative action law had a significant net effect.

Next, we chronicle the evolving treatment of pregnancy under sex discrimination law, developing hypotheses about the effects of three different

policy approaches. We then review competing theories of employer maternity leave provision before turning to our central analysis.

## THE LAW AND MATERNITY LEAVE

We examine the effects of three laws to make two critical comparisons. First, we examine two substantively similar laws: a contested administrative ruling and an uncontested congressional act. We compare employer response to the EEOC guidelines of 1972 and the Pregnancy Discrimination Act of 1978. Second, we examine two substantively different laws: an equal treatment law requiring employers to match maternity benefits to disability benefits and an affirmative action law requiring all employers to offer maternity leave. We compare employers subject to the Pregnancy Discrimination Act of 1978 with those additionally subject to California's 1978 amendments to its Fair Housing and Employment law.

### Before Maternity Leave Laws

Before the early 1970s, the American state did little to promote maternity leave (Kamerman, Kahn, and Kingston 1983). During World War II, the Department of Labor's Women's Bureau recommended that employers guarantee the jobs of women who were absent due to pregnancy and maternity. But most employers continued to require pregnant women to resign from their jobs (Frank and Lipner 1988, p. 15; Silverman 1943). The Equal Pay Act of 1963 required employers to pay men and women the same wages for the same work, and the Civil Rights Act of 1964 outlawed employment discrimination based on sex. However, neither act mentioned maternity or pregnancy. In 1966, the Equal Employment Opportunity Commission (EEOC), created to administer the employment provisions (Title VII) of the Civil Rights Act, stated that providing "truly equal employment opportunities" meant guaranteeing the jobs of women who were absent from work due to maternity and childbirth (EEOC 1966, p. 40). Yet the EEOC also told employers that it was not illegal to exclude maternity from disability leave programs (Vogel 1993, pp. 64–65).

   Employers felt little pressure to offer maternity leave in the years after the passage of the Civil Rights Act. A 1967 study of large employers, conducted by the Bureau of National Affairs (BNA), found that 18% still required pregnant employees to resign (BNA 1967, p. 20). The vice president of the First National Bank of Cincinnati pointed to the EEOC's ambivalence in explaining his company's resignation policy: "It is my understanding that while the EEOC favors the concept of maternity leaves, it has not agreed to a policy statement. Until such time, I shall follow my

former policy" (quoted in BNA 1967, p. 20). During the 1960s, civil rights legislation did little to encourage employers to offer maternity leave.

### The 1972 EEOC Ruling: Promulgation by Litigation

In March of 1972, the EEOC issued new Guidelines on Discrimination because of Sex that defined pregnancy discrimination as an element of sex discrimination. Employers could not refuse to hire, train, or promote a woman because of pregnancy (29 Code of Federal Regulations 1604 [1972]). Neither could they fire a woman because of pregnancy. The guidelines also directed employers to treat pregnancy-related disabilities as they treated other disabilities that keep workers away from their jobs for short periods. Employers who provided leaves, health insurance, or income replacement for nonoccupational injury and illness were required to do the same for pregnancy-related disabilities. The EEOC guidelines also suggested that employers who failed to allow leaves for pregnancy-related disability were guilty of discrimination, unless they could prove a business necessity for forbidding leaves.

Using the EEOC guidelines, the EEOC and unions sued employers over their maternity leave benefits. Employers fought back by charging that the EEOC guidelines were an illegal deviation from Congress's intent in the Civil Rights Act of 1964. The EEOC's targeted lawsuits produced a great deal of publicity for the new maternity leave regulations. The guidelines were effective, paradoxically, because as administrative law they were vulnerable to reversal by the courts; when employers fought the new regulations, they helped generate even more publicity for the issue of maternity leave. The EEOC apparently recognized this when they initiated suits against high-profile employers, seeking both a favorable Supreme Court ruling and publicity for the guidelines. While the Supreme Court overturned the guidelines in 1976, many employers who followed the legal contest did not wait for the standing of the guidelines to be clarified.

*The women's movement and the EEOC regulations.*—We show below that after the EEOC ruling, employers were more likely to create maternity leave programs. Some might argue that the law was merely a mediating force and that the women's movement pressured the EEOC and employers to change their maternity leave policies. Did the women's movement prompt the EEOC ruling and employers' adoption of maternity leave? We now take a brief detour to answer this question. First, we note that scholars of the women's movement give feminists within the federal government a central role in stimulating the women's movement (see Harrison 1988; Costain 1992). Kennedy's Commission on the Status of Women gave rise to state-level commissions, and it was at a 1966 con-

ference of these commissions that feminists encouraged Betty Friedan to found the National Organization for Women (NOW) as an external lobbying body (Harrison 1988, chap. 9; Costain 1992, pp. 44–45; Skrentny, in press). Second, by the early 1970s, NOW and other women's movements groups were functioning autonomously, but they were not concentrating on federal maternity leave policy. At this time, NOW was busy lobbying states for passage of the Equal Rights Amendment and lobbying the EEOC to end discrimination in job advertising.

Instead, it was the federal Citizens' Advisory Council on the Status of Women (successor to Kennedy's Commission) that led the charge to change federal policy on maternity leave. The Council developed sample maternity leave policies for federal employees, pressed the EEOC for a ruling on maternity leave, and provided the language for the 1972 maternity leave guidelines (Citizens' Advisory Council on the Status of Women 1971, 1973; Vogel 1993, pp. 62–65). The historical evidence suggests that feminists *within* the federal government lobbied for the maternity leave ruling, which later led employers to change their practices.

*The EEOC guidelines and employer response.*—Contemporaneous studies showed that large employers responded quickly to the new EEOC regulations. A Prentice Hall study of 108 large employers found that more than half were at least considering new maternity leave policies in the wake of the EEOC's guidelines (Hyatt 1972, p. 20). EEOC Commissioner Ethel Bent Walsh reported in 1974 in the *Labor Law Journal* that over three-fourths of banks and insurance companies had liberalized their maternity leave policies and benefits "as a direct result of [EEOC] Guidelines" (Walsh 1974, pp. 153–54). One study found that between 1972 and 1975, two-fifths of employers made major changes in their maternity leave policies (BNA 1975). Another found that the number of firms offering maternity leave nearly tripled between 1969 and 1978 (Kamerman et al. 1983, p. 56). We hypothesize that employers responded to the EEOC's new position with new maternity leave policies.

HYPOTHESIS 1.—*In the wake of the EEOC guidelines, employers were significantly more likely to create maternity leave policies.*

*Legal challenges to the EEOC ruling.*—After the EEOC issued the 1972 guidelines, the Commission began a legal campaign against large, high-profile employers. Some of these employers, notably General Electric, challenged the EEOC's interpretation of Title VII. The G.E. case (*General Electric Co. v. Gilbert et al.,* 97 Sup. Ct. 401 [1976]) dealt directly with disability insurance (i.e., income replacement) during leaves but drew attention to the wider issue of maternity leave. Using the EEOC guidelines, the United Electrical Workers union initiated a class action suit, in which they claimed that G.E.'s disability benefits program discriminated against

pregnant women. General Electric argued that the EEOC's interpretation of Title VII as covering pregnancy ran counter to the legislative history of the Civil Rights Act.

Before the Supreme Court's 1976 ruling in favor of G.E., both the District Court of Virginia and the United States Court of Appeals for the Fourth Circuit ruled that G.E. had discriminated against pregnant employees (Fryburger 1975). These decisions created modest financial sanctions and tremendous negative publicity for the company. In April of 1974, district court judge Robert Merhige ruled that G.E. must treat pregnancy as it treats other temporary disabilities: "To isolate such a disability for less favorable treatment in a scheme purportedly designed to relieve the economic burden of physical incapacity is discrimination by sex" (quoted in *National Underwriter* 1974, p. 4). G.E. had argued that because pregnancy is voluntary, it is not like other disabilities. The court found that because G.E. does not exclude any other disabilities, they cannot exclude pregnancy-related disabilities: "That this is sex discrimination is self-evident" (quoted in *National Underwriter* 1974, p. 4). The court of appeals upheld this decision in 1976.

Late in 1976, the Supreme Court reversed the lower court decisions in the G.E. case, ruling that employers who treat pregnancy differently from other nonoccupational disabilities are *not* guilty of sex discrimination. Justice Rehnquist, for the Court, argued that excluding a condition from disability leave and benefit coverage as "a subterfuge to accomplish a forbidden discrimination," in this case against women, would be illegal if the condition were similar to covered conditions. But pregnancy is "not a 'disease' at all, and is often a voluntarily undertaken and desired condition" (*General Electric Co. v. Gilbert et al.,* 9). The Court rejected Judge Merhige's position that maternity should be treated like cosmetic surgery, which is voluntarily undertaken, and like sports injuries, which occur during recreational activities that carry known risks. The G.E. challenge seemed to have succeeded in the courts, but the notoriety it brought encouraged employers to add maternity leave policies.

*Press coverage as the mediating force.*—The EEOC's announcement of the guidelines produced a brief burst of press attention, but legal challenges generated the ongoing publicity that brought the EEOC's position to the attention of most employers. The 1972 guidelines made the first page of the *New York Times* (March 31, 1972), the sixth page of the *Wall Street Journal* (April 3, 1972), and the tenth page of section three of the *Los Angeles Times* (April 1, 1972). The *New York Times* reported that, "In its most sweeping revision of its guidelines on discrimination because of sex since they were first adopted in 1966, the employment commission now says that to deny a woman a job because she is pregnant is, on its face, a violation of the Civil Rights Act of 1964.... In addition, disabilities

related to pregnancy, including recovery from childbirth, miscarriage, and abortion, should be treated by employers the same as any other temporary disability in terms of leave" (*New York Times* 1972, p. A1).

In April of 1974, the *New York Times* reported that in the two years since the EEOC's guidelines were issued, the commission had sued 11 large employers for maternity discrimination, with 111,052-employee du Pont at the top of the list. In reporting on these suits, the *Times* warned, "Just as women have always know that it is impossible to be 'a little bit pregnant,' so the nation's employers are learning—from the Federal Government—that it is imprudent to be a little bit prejudiced against a pregnant employee" (Dullea 1974, p. 41). Specialist journals also advised employers to add maternity leave. Harry Edwards argued in the *Labor Law Journal* in 1973 (p. 421), "reinstatements after childbirth will probably be required unless the employer can show that there is some compelling 'business necessity' which justifies a refusal to reinstate." *Business Week* reported that after the district court's decision in the G.E. case in April of 1974, the EEOC filed 130 pregnancy discrimination suits and issued warnings to another 250 companies. The union that fought the case against G.E. soon brought charges against another 73 employers, and *Business Week* (1976, p. 41) noted that "Companies on the firing line include Westinghouse, RCA, Chrysler, Sperry Rand, Pennwalt, Honeywell, Philco-Ford, and Phelps Dodge." Given the prominence of the EEOC guidelines and pursuant lawsuits at the time, it is striking that a quarter-century hence, politicians, managers, and scholars have forgotten the role of the state in popularizing maternity leave.

The G.E. case and related litigation received much more press coverage than the original guidelines had. While the adoption of the guidelines merited one or two articles in each major daily paper we examined, the G.E. case was the subject of 20 *New York Times* articles between 1972 and 1976 and was covered widely in the pages of *Business Week, Fortune,* and *Forbes.* The district court and court of appeals decisions of 1974 and 1976, backing the guidelines, received particularly extensive coverage. Figure 1, which charts *New York Times* articles on maternity leave issues from 1970 to 1984, illustrates this. There was a dramatic rise in the number of articles in 1972, the year that saw the issuance of the EEOC guidelines and the initiation of the G.E. case. The number of articles remained high for the course of the litigation.

We hypothesize that early lawsuits generated publicity about the new regulations and so led employers to install maternity leave programs, despite the fact that the lawsuits challenged the validity of the EEOC ruling. Press coverage mediated the effects of lawsuits. Figure 2 shows that in our sample, the proportion of firms with maternity leave policies rose dramatically between 1974 and 1977. Compare this with figure 1, which
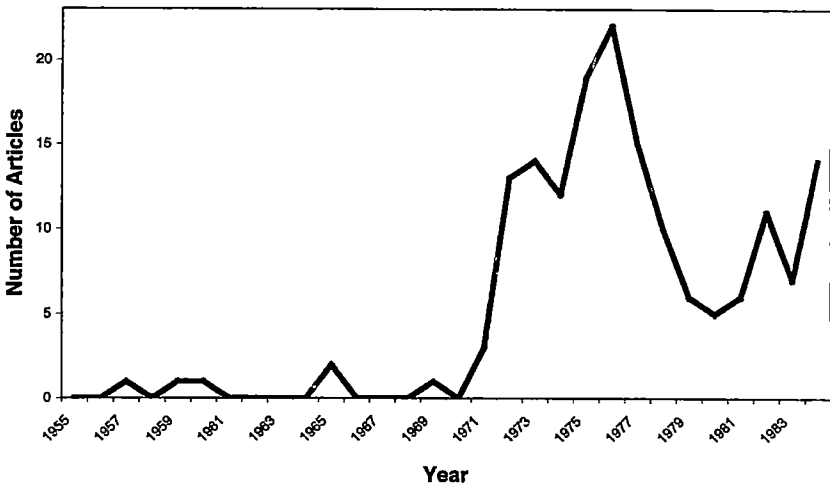
FIG. 1.—*New York Times* articles on maternity leave

shows press coverage of maternity leave. While it is important to exercise caution in interpreting figure 2, because changes over time may simply result from changes in covariates such as size, the correspondence between figures 1 and 2 during this period is suggestive. It appears that press coverage stimulated employers to adopt maternity leave. There were few suits
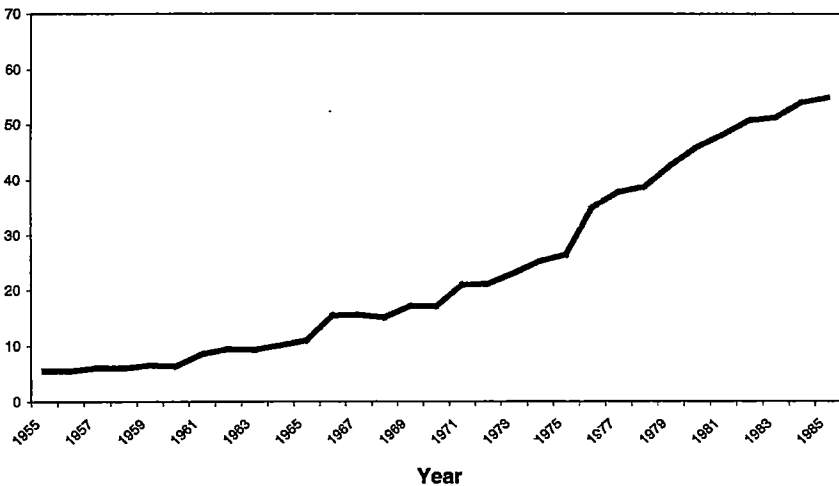


FIG. 2.—Percentage of employers with maternity leave

in these early years, thus, as we argue below, employers responded to publicity rather than to the actual risk of being sued.

HYPOTHESIS 2.—*Press coverage of maternity leave regulations increases the likelihood an employer will adopt maternity leave policies.*

The Pregnancy Discrimination Act of 1978: Congress's Equal Protection Mandate

After the Supreme Court struck down the EEOC ruling in 1976, Congress responded by writing the EEOC position into law in the Pregnancy Discrimination Act (PDA) of 1978. The PDA gives us the opportunity to compare the effects of an administrative ruling with those of a similar congressional act. The PDA amended the Title VII definition of "discrimination because of sex" to include discrimination "on the basis of pregnancy, childbirth, or related medical conditions." Like the EEOC ruling, the act depended on the principle of equal protection to require employers to treat women affected by pregnancy "the same for all employment-related purposes" as other persons "similar in their ability or inability to work" (United States Code 42: 2000e[k] [1994]). Employers would have to extend disability leaves and pay to cover pregnant women.

Employers did not challenge the PDA in court as they had challenged the EEOC guidelines, and as a consequence, the PDA received little press coverage. G.E. had argued that the EEOC's 1972 ruling misinterpreted congressional intent. Employers had no grounds to challenge the 1978 statute because Congress had made its intent clear therein. After a brief burst of press coverage, the PDA received little attention. While there had been 109 *New York Times* articles on maternity leave between 1972 and 1978, there were just 43 between 1979 and 1984 (see fig. 1). Most of those focused on new state and local maternity leave regulations.

Before 1978, the legal standing of the EEOC ruling was uncertain. Several courts found that Title VII of the Civil Rights Act required employers with disability benefits to offer matching maternity benefits. But G.E. continued to fight the case, and the Supreme Court found that Title VII required no such thing. Students of law and regulation predict that when a ruling is challenged in the courts, firms will delay compliance and await a resolution. If employers had awaited a final decision, then we would expect to see a slow increase in maternity leave policies after the EEOC ruling, followed by a rapid increase after the PDA. Instead, we expect that wide press coverage of G.E.'s challenge of the EEOC ruling led many employers to create maternity leave programs before the PDA settled the legal issue. Under our scenario, adoption rates would be high in the mid-1970s. The Pregnancy Discrimination Act may have led firms to liberalize existing maternity leave programs (Vogel 1993) and encouraged feminists

to fight for broader family leave laws (Elving 1995), but we expect that the law spawned few new corporate maternity leave programs. In resolving the legal issue, the PDA had the effect of curtailing public discourse on maternity leave.

It is important to note that studies of the spread of other corporate equal opportunity practices *do not* find a slowdown during the 1980s, after the Reagan administration reduced enforcement of Civil Rights laws. Employers continued to adopt affirmative action offices (Edelman 1992; Dobbin and Sutton 1998), formal grievance procedures to intercept discrimination complaints (Sutton et al. 1994; Sutton and Dobbin 1996), and formal hiring and promotion practices to prevent discrimination by middle managers (Dobbin et al. 1993). Yet we predict that publicity surrounding the G.E. case had led most employers who monitor the legal environment to install maternity leave policies by the late 1970s. A contested administrative ruling had caused susceptible employers to change their practices, so the new statute had limited effects.

HYPOTHESIS 3.—*Because many employers established maternity leave programs following the EEOC guidelines of 1972, we expect that the Pregnancy Discrimination Act will have little effect on employer adoption of maternity leave programs.*

## California's Maternity Leave Legislation of 1978: The Affirmative Action Approach

Some states went a step further than Congress, requiring all employers to offer job-protected maternity leave. California, in a 1978 amendment to the state Fair Employment and Housing Act (California Government Code 12945 [1998]), required employers with 15 or more workers to allow women to take maternity leaves. Unlike the PDA, the California law did not state that maternity leave provisions should match disability leave benefits. Instead it required all but the smallest of employers to allow job-guaranteed maternity disability leaves of up to four months, regardless of their treatment of other disabilities. For example, even if an employer only held a worker's job for four weeks after a heart attack, the employer would be required to hold a woman's job for 16 weeks if she was unable to work because of pregnancy or recovery from childbirth. Whereas the PDA was based on the principle of equal opportunity, the California law was based on the principle of affirmative action—it was designed to improve opportunities for a particular protected group.

While the issue of affirmative action in employment has generated a great deal of heat, scholarship has as yet generated little light on its relative efficacy. Previous labor market studies have compared sectors subject to equal opportunity law to those additionally subject to affirmative action

law (Ashenfelter and Heckman 1976; Leonard 1984). Black men made bigger inroads in the latter group than in the former. We employ a similar method to examine the effects of these two approaches on employer practices.

We hypothesize that California's affirmative action law was more effective than the federal equal treatment law. History has produced a straightforward test. By 1978, the PDA required that maternity leaves and pay match disability leaves and pay. This law applied to employers in all 50 states. The California employers in our sample were, additionally, required by that state's affirmative action law to offer maternity leaves. Thus, in comparing California employers to those in New Jersey and Virginia after the 1978 legislation, we have a clear test of the effect of adding an affirmative action requirement to an equal treatment requirement.

Like the EEOC regulations, the California law received some press coverage upon passage and more attention during the course of legal challenges. A *Los Angeles Times* editorial encouraged passage in August of 1978, and an article in the same paper announced passage on September 30, 1978. Employers challenged the law in 1982, claiming that it conflicted with the equal treatment principle of the Pregnancy Discrimination Act (*California Federal Savings and Loan Association v. Guerra*, 107 Sup. Ct. 683 [1987]). In the 1987 Supreme Court decision upholding the California law, Thurgood Marshall argued that the Pregnancy Discrimination Act was intended to create "a floor beneath which pregnancy disability benefits may not drop—not a ceiling above which they may not rise" (quoted in Vogel 1993, p. 88; see also Rhode 1989). The case received considerable national coverage; five of the *New York Times* articles on maternity between 1982 and 1984 (see fig. 1) covered it.

Feminists held different positions on the wisdom of the affirmative action approach, and the public debate helped to make employers aware of the California law (Gorney 1984; Rhode 1989). Some feminists opposed what they derogatorily referred to as the "special treatment" approach of the California law, drawing comparisons to paternalist legislation from the turn of the century that had limited women's employment opportunities (see Kessler-Harris 1982; McCammon 1996). Many mainstream organizations that had fought for gender equality—the National Organization of Women, the League of Women Voters, the National Women's Political Caucus, and the American Civil Liberties Union—sided with California employers in opposing that state's law (Vogel 1993, p. 78). These groups advocated the approach that would later become law in the Family and Medical Leave Act, of guaranteeing leaves to new fathers as well as new mothers. When the United States District Court overturned California's maternity leave law in 1984, Dianne Feinstein, San Francisco's Democratic mayor, praised the court, arguing, "What we women have been
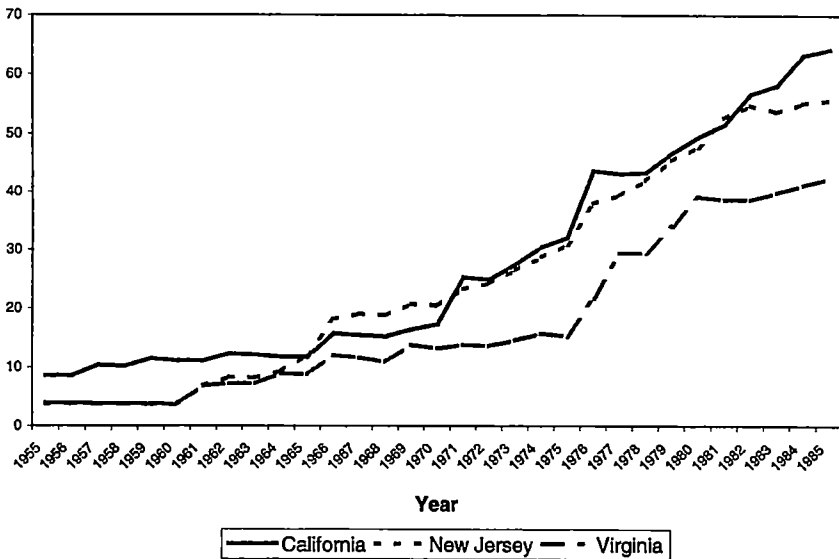
FIG. 3.—Percentage of employers with maternity leave by state

saying all along is we want to be treated equally. Now we have to put our money where our mouth is. What we were asking was to create a special group of workers that in essence is pregnant women and new mothers. I just don't happen to agree with that" (UPI 1984, p. 5). Others argued that because women bear the brunt of pregnancy, requiring equal treatment did not do enough to equalize opportunity for women. These feminists saw California's law as a more realistic way to help advance women's employment opportunities.

We predict that California's affirmative action law had a net positive effect on employers. In figure 3, we present the prevalence of maternity leave, both over time and by state. By about 1980, the diffusion of maternity leave policies had stalled among New Jersey and Virginia employers but not among California employers. Others have found that state-level employment statutes do shape workplace practices (e.g., Guthrie and Roth 1999).

HYPOTHESIS 4.—*Following the maternity leave requirements of the 1978 law, employers in California will adopt maternity leave at higher rates than employers in New Jersey or Virginia.*

THEORIES OF FRINGE BENEFITS

The conventional wisdom, as sketched above by Bob Dole and the Chamber of Commerce, comes from labor economists, who suggest that employ-

ers responded to the growing ranks of women in the labor market with maternity leave programs. We examine this thesis by looking at industry feminization and change in feminization. A second important argument comes from students of law and regulation, who suggest that employers comply with laws in direct proportion to the risk of sanction. We examine this thesis by comparing a congressional statute of sure legal standing with an administrative ruling that was being challenged in the courts. We also examine the risk of legal sanction with a measure of maternity leave cases in federal courts.

Quite a few scholars have studied employer work/family policies in general and maternity leave in particular, although none has studied the spread of maternity leave over time (Glass and Fujimoto 1995; Goodstein 1994; Guthrie and Roth 1999; Ingram and Simons 1995; Milliken et al. 1998; Osterman 1995). Drawing on these studies, we review other factors that have affected human resources practices and work/family policies in particular.

### The Feminization Thesis

Labor economists argue that employers use benefits to compete for labor and tailor benefits to labor market characteristics. A number of analysts have sought to show that industry feminization predicts employer work/family benefits. Cross-sectional studies have produced mixed results. Goodstein (1994) and Osterman (1995) find that employers with large numbers of women are more likely to offer work/family benefits, and Guthrie and Roth (1999) find that employers in industries with large numbers of women are more likely to offer paid maternity leave. But Glass and Fujimoto (1995) found that employers with large numbers of women are not more likely to have maternity leave, and Ingram and Simons (1995) find that such employers are no more likely to offer an array of work/family benefits, including maternity leave. Previous cross-sectional studies have not examined whether trends in feminization matter. If labor economists are correct, we should find that industry feminization predicts maternity leave adoption and that change in feminization also predicts adoption.

HYPOTHESIS 5.—*Employers in sectors with high proportions of women workers, and those in sectors where women's representation is growing, will adopt maternity leave to compete in the labor market.*

### The Risk of Sanction Thesis

Students of regulation from both the left and right have made economistic arguments about how firms respond to the law. Richard Posner, a founder

and prolific proponent of the school of law and economics, suggests that firms are rational actors when dealing with the law (Posner 1974, 1997). When faced with a law, they calculate the costs of compliance and non-compliance. As the risk of legal sanction increases, so should the calculated cost of noncompliance. Some students of regulation from the left, such as Cass Sunstein (1990, 1996), advocate a similar approach to the issue of compliance as well. Agents comply with the law in proportion to the perceived risk of sanction.

Based in economic arguments about behavior, these perspectives suggest hypotheses about employer compliance with maternity leave law. First, they suggest that employers should have been more likely to create maternity leave programs after passage of the PDA, which had certain legal standing, than after the EEOC ruling, which had uncertain legal standing and was in fact struck down by the Supreme Court.

HYPOTHESIS 6.—*Because the legal standing of the EEOC guidelines was uncertain and the standing of the PDA was not, employers awaited the latter before creating maternity leave policies.*

Second, employers should have adopted maternity leave in direct response to the rising risk of being sued. Figure 4, which charts the number of federal court decisions in maternity leave cases, shows the increased risk of litigation faced by employers. While a few high-profile lawsuits dominated the media in the early and mid-1970s, the number of suits was small. Litigation increased after press coverage rose and remained high
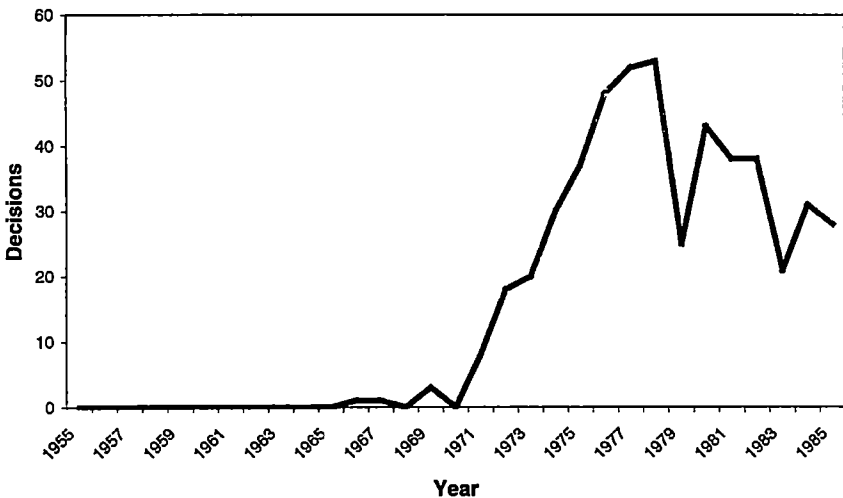


FIG. 4.—Federal court decisions in maternity leave cases

after the Pregnancy Discrimination Act and during the Reagan era. In his study of pregnancy discrimination in the United States, Mark Edwards argues that this litigation led employers to change their maternity leave practices. He claims litigation "directly forced employers to accommodate pregnancy in the workplace" (1996, p. 251). We examine whether there is a direct effect of litigation. Previous studies suggest that the risk of litigation affected employer maternity leave practices even after passage of the FMLA. In a study of corporate maternity pay in 1996, Guthrie and Roth (1999) show that employers in liberal federal circuit court districts are more likely to offer pay. We test the hypothesis that the risk of litigation stimulated the spread of maternity leave.

HYPOTHESIS 7.—*The risk of litigation, as measure by an annual count of federal court decisions in maternity leave cases, increases the likelihood an employer will adopt maternity leave.*

Neoclassical and Transaction Cost Economics: Recruitment and Retention

Economists offer several relevant hypotheses, beyond the feminization thesis, that concern how employers use fringe benefits to attract and retain workers. Firms are thought to provide better benefits when labor markets are tight and when turnover is particularly costly.

There is some evidence that employers are more likely to adopt maternity leave and other family benefits in tight labor markets. Glass and Fujimoto (1995) find that employers with large numbers of professional and managerial workers—two groups with low unemployment rates—are more likely to offer maternity pay but no more likely to offer job-guaranteed leaves. Goodstein (1994) and Ingram and Simons (1995) find that industry-level scarcity of female workers leads employers to offer work/family benefits such as maternity leave, but Osterman (1995) finds that employers with recruitment and retention problems are no more likely to have work/family policies.

HYPOTHESIS 8.—*Employers in sectors with low unemployment rates, and in sectors with declining unemployment rates, will be more likely to adopt maternity leave.*

Labor economists and transaction costs theorists expect employers who depend on firm-specific skills, such as those in capital-intensive industries, to offer benefits that bind workers to the firm (Doeringer and Piore 1971; Williamson 1981). Sociologists and economists have found that firms in the finance sector also depend on long-term employment, and these employers also use personnel practices and fringe benefits to tie workers to the firm (Goldin 1986; Baron, Jennings, and Dobbin 1988; Milliken et al. 1998).

HYPOTHESIS 9.—*Industries that value long-term employment will be more likely to adopt maternity leave policies.*

Labor segmentation theorists have made kindred arguments about employer efforts to retain workers through "bureaucratic control" systems, which formalize promotion and expand retention-related benefits (e.g., pensions and leaves) to encourage workers to stay with the firm (Edwards 1979; Gordon, Edwards, and Reich 1982). Guthrie and Roth (1999) find that employers with formalized personnel systems are more likely to offer paid maternity leave.

HYPOTHESIS 10.—*Employers with formal internal labor market practices will be more likely to adopt maternity leave.*

### Neoinstitutional Theory: Professional Groups and Public Visibility

Neoinstitutional studies have established that professionals often determine corporate response to unclear policy mandates (Edelman et al. 1992; Sutton and Dobbin 1996, p. 795). Personnel and legal experts have promoted a range of different Civil Rights compliance mechanisms. Benefits specialists were in charge of leave policies, and so they were most likely to have monitored legal developments in this area.

HYPOTHESIS 11.—*Organizations with personnel, legal, and benefits offices will be more likely to adopt maternity leave policies.*

Studies have also shown that public sector and nonprofit organizations are more susceptible to normative pressures regarding fair employment practices (Meyer and Scott 1983; Dobbin et al. 1988; Edelman 1990). In an analysis of cross-sectional data from 1991, Ingram and Simons (1995) confirm that public-sector employers are more likely to offer work/family benefits, including leaves, dependent care assistance, and flexible work schedules.

HYPOTHESIS 12.—*Public and nonprofit organizations will be more likely than private sector firms to adopt maternity leave policies.*

### Industrial Relations: Union Pressure

Unions have historically fought for generous fringe benefit packages (Freeman 1981; Freeman and Kleiner 1990; Kalleberg and Van Buren 1996), but studies find little evidence that unionization is associated with leave policies (Osterman 1995; BNA 1986). We control for unionization.

HYPOTHESIS 13.—*The presence of a union will increase an employer's likelihood of adopting maternity leave.*

Unions in feminized sectors have often pushed for benefits favored by women (Milkman 1987, p. 2; Schatz 1983). Glass and Fujimoto (1995) find that unionization *and* an interaction term between unionization and

feminization show positive effects on organizational maternity leave, suggesting that unions fight for maternity leave when their members are women.

HYPOTHESIS 14.—*Unionization will have a particularly strong effect in feminized industries.*

## Size and Age

Scale should predict the formalization of maternity policies, as it predicts the formalization of many other policies (Blau and Schoenherr 1971; Kalleberg and Van Buren 1996). Most studies show that size has a nonlinear effect, because an increase of 100 employees has a greater impact on a firm of 50 than on a firm of 1,000.

HYPOTHESIS 15.—*Size will have a positive, but declining, effect on the likelihood employers will adopt maternity leave policies.*

Older organizations should be less likely to adopt maternity leave. Stinchcombe (1965) discovers that organizations find it difficult to change established practices. Selznick (1957) likewise argues that, over time, organizations become resistant to change.

HYPOTHESIS 16.—*Older organizations will be less likely to adopt maternity leave.*

## DATA AND METHODS

### Sample and Data Collection

We analyze data collected from a stratified random sample of public, nonprofit, and private sector establishments in 1985–86. Retrospective data on the history of each establishment's personnel policies were collected from 279 respondents. These respondents were located in three states, selected for their varying legal environments: California, New Jersey, and Virginia. The sample was stratified by industry to ensure variety and to facilitate examination of industry effects. Establishments were sampled from banking, chemical manufacturing, electrical manufacturing, machinery manufacturing, publishing, transportation, retail trade, hospitals, nonprofit organizations, and city, county, state, and federal agencies. We used the best population list available for each group and omitted employers with fewer than 50 workers and employers who reported that they did not operate in one of the states or in one of the industries sampled (Dobbin et al. 1993; Sutton and Dobbin 1996). The survey covered organizational structure and demographics and key employment policies for the period 1955–85. We obtained a response rate of 48%, but after eliminating 21 responses due to the poor quality of the data, we were left with a net

rate of 45%. This compares favorably with the response rates obtained in other organizational studies. Milliken et al. (1998) report a rate of 18%; Lincoln and Kalleberg (1985) report 35%; Blau et al. (1976) report 36%; Edelman (1992) reports 54%; and the National Organizations Survey for 1991 reports 65% (see Spaeth and O'Rourke 1996, p. 32). We expect that there is some response bias in the data, in that establishments with personnel offices were probably more likely to participate. As such establishments may be more likely to create maternity leave policies, we suspect that our over-time plots somewhat overstate the prevalence of maternity leave. This bias is unlikely to affect the event history analyses, however. First, the survey includes information on the factor most likely to create response bias, the presence of a personnel office, and, hence, we can include that factor in the model. Second, we have many organization-by-year spells without personnel offices, and so we have adequate variance on this covariate.

The data were transformed into annual spells, yielding a data set with 4,921 at-risk organization-years. Each annual spell records values for the outcome and for independent variables for each year. We retained a record for each year in which an organization was operating and at risk of maternity leave adoption. We removed the organization-year spells in which each organization was not at risk of adopting maternity leave—those spells before the organization's birth and those after it had created a leave policy. Organizations that did not adopt maternity leave before 1986 are included for the whole period. Our analysis relies on the 4,726 at-risk organization-years with complete data for the relevant variables.

### Measures and Model Specification

*Maternity leave policy adoption.*—Survey respondents were asked whether their organization had ever had a formal policy covering maternity leave. As noted above, organizations remain in the risk set until adoption of maternity leave or censoring, in 1985. We do not have information on whether the leave policies provide pay or on the length of leave. For organizations with policies, respondents were asked when the policy was adopted and when, if ever, it was abandoned. Formal policies covering maternity leave serve as an important resource for women. For the most part, the policies adopted in the 1970s and early 1980s replaced formal or informal policies of requiring pregnant workers to resign (Fryburger 1975; O'Connell 1990). The job guarantee fosters career continuity, and this has positive long-term effects on women's advancement and income (Waldfogel 1997; Jacobsen and Levin 1995).

*Independent variables.*—Table 1 lists the independent variables. These vary over time, and they were collected at the establishment level unless

## TABLE 1

### VARIABLES USED IN ANALYSIS

| Variable | Description | Predicted Effect |
|---|---|---|
| Organizational characteristics: | | |
| Log employment | Natural logarithm of number of employees in previous year | + |
| Age | Years since organization's founding | − |
| Banking | Binary variable marking sector—does not change over time | + |
| Core manufacturing | Chemicals, electrical manufacturing, and machinery manufacturing | + |
| Government agencies | Federal, state, county, and local government | + |
| Unionization | Binary variable for presence of union contract in previous year | + |
| Personnel office | Binary variable for presence of personnel office in previous year | + |
| Benefits office | Binary variable for presence of benefits office in previous year | + |
| Legal office | Binary variable for presence of legal office in previous year | + |
| Formalization scale | Scale (1–5) marking presence of job ladders, written job descriptions, salary classification system, nonunion grievance procedure, and performance evaluations in previous year | + |
| Labor market factors: | | |
| Unemployment | Regional unemployment rate in previous year | − |
| Change in unemployment | Change in regional unemployment | − |
| Feminization | Industry feminization in previous year | + |
| Change in feminization | Change in industry feminization | + |
| Legal and press attention: | | |
| Litigation | Federal decisions in maternity leave cases (three-year moving average) | + |
| Press coverage | *New York Times* articles on maternity leave (three-year moving average) | + |
| Policy environment: | | |
| 1955–73 | Reference period | |
| 1974–79 | Period following EEOC guidelines | + |
| 1980–85 | Period following Pregnancy Discrimination Act | (none) |
| California | Reference states are New Jersey and Virginia | (none) |
| 1974–79 | Additional effects of period following EEOC guidelines, in California | (none) |
| 1980–85 | Period following California's affirmative action law | + |

NOTE.—Variables are subject to change over time. Reference industries include nonprofit organizations, hospitals, publishing, transportation, and real estate

noted. To measure establishment size, we requested figures on the number of employees in 1985, 1975, 1965, and 1955. We also asked about years in which unusual gains or losses occurred. Annual size figures were interpolated from these responses and logged for use in the models. To measure age, we collected information on the founding date of each organization. We measure firm-specific skills with industry variables. Previous studies show that among the industries in our sample, chemicals, electrical machinery, and machinery manufacturing have high levels of firm-specific skills (Gordon, Edwards, and Reich 1982). We include these three industries in core manufacturing. We also include banking and government variables to test the hypotheses that concern for long-term employment and vulnerability to state and public pressures increase adoptions of maternity leave. We use a binary variable that represents the presence of a union contract. To measure formalization, we include a scale comprised of five internal labor market (ILM) practices: job ladder, written job description, salary classification, nonunion grievance procedure, and formal performance evaluation. We measure the presence of legal, personnel, and benefits offices with binary variables.

We supplement the survey data with publicly available data on labor market characteristics. To measure the proportion female in each industry, we use data from the Bureau of Labor Statistics' Current Employment Statistics (CES) Survey, which provides annual figures on female employment at the two-digit Standard Industrial Code (SIC) level. CES is the only source of labor force data that covers feminization before 1964, and it provides a more detailed industry breakdown than the other possible sources, such as the early Current Population Survey and the Equal Employment Opportunity Commission annual reports. For a few industries, we estimated figures for several missing years using simple linear regression.[3] Most missing data were from 1955–57, when there was little annual change in female labor force participation. In addition to proportion female, we use a lagged change score. While we would like to include an establishment-level measure of feminization as well, questions on workforce composition were deleted from the survey instrument after a pretest showed a high rate of refusals for questions about female employment.

To measure unemployment, we use regional unemployment rates calculated from the Current Population Survey. We use regional data rather

---

[3] The estimated years are 1955–57 (chemicals, publishing); 1955–58 (electrical, machinery, hospitals, retail trade); 1955–62 (banks, transportation, government); and 1955–71 (nonprofit organizations).

than state-level data, because these data are available for the entire period. We use both the unemployment rate and a change score.

To measure press attention, we use a lagged three-year average of the number of *New York Times* articles dealing with maternity leave. The *New York Times* is cataloged in the *New York Times Index* and in the Lexis-Nexis service. These sources made identifying and counting maternity leave articles feasible, whereas other papers and magazines are not well indexed for the entire period covered by our study. We also estimated models using a lagged annual count of articles, and the results were substantially similar.

To measure litigation activity, we use a lagged three-year average of the number of federal court decisions in cases dealing with maternity leave. We gathered annual count data from the Lexis legal research service, using the GENFED, NEWER libraries to search all federal cases since 1944. We also collected annual counts of decisions in cases related to pregnancy discrimination. We estimated models using a lagged count of maternity leave cases, a lagged count of maternity leave and pregnancy discrimination cases, a lagged count of pregnancy discrimination cases, and a three-year average of pregnancy discrimination cases. Results were substantially similar, hence, we report models that use the lagged three-year average of maternity leave cases.

To measure the effects of public policies, we include binary variables for the periods 1974–79 and 1980–85 (with 1955–73 omitted), a binary variable for operation in California, and an interaction between California and each period variable. We lag the periods one year, beginning each in the January following the first full year of operation of the new policy. We expect that the contentious legal environment, prompted by changes in the EEOC guidelines, will increase adoption among all employers between 1974 and 1979. We also hypothesize that the federal Pregnancy Discrimination Act, which required equal treatment of pregnancy and other disabilities, will not increase the adoption of formal maternity leave policies in New Jersey and Virginia in the period 1980–85. We expect that California's maternity leave requirement, added to its state fair employment law late in 1978, will increase adoptions in that state between 1980 and 1985. Thus, we expect that, with 1955–73 as the reference period, the period 1974–79 will have a positive effect, and the period 1980–85 will have little or no effect. We expect to find that California was no different from New Jersey and Virginia before 1980 and thus do not expect to see significant coefficients for California or for California, 1974–79. However, we expect the interaction of California with 1980–85 to have a positive effect as a consequence of the state's maternity leave law.

Estimation

We use discrete event-history methods because we do not know the exact timing of adoption within the spells and because we have many "tied" events, that is, years in which multiple employers adopt maternity leaves (Kalbfleisch and Prentice 1980; Allison 1995, pp. 220–22). In particular, we employ discrete models where the hazard (instantaneous risk) at time $t$ for an organization with characteristics $i$ is

$$h(t|X_i) = h_0(t)^{(X_i'\beta)}.$$

Here, $h_0(t)$ is a baseline hazard function describing the risk for organizations with baseline characteristics $X = 0$, and the exponentiate $(X_i' \beta)$ is a proportionate increase or reduction in risk associated with characteristics $X_i$.

Because the transformed data set contains annual spells, the hazard of adoption in each year is equivalent to $P_{it}$, the conditional probability that the event occurs to organization $i$ in year $t$, given that it has not already occurred. We model the complementary log-log transformation of the survival function $(1 - P_{it})$ on the covariates, specified for that time:[4]

$$\log[-\log(1 - P_{it})] = \alpha(t) + \beta'X_{it}.$$

The coefficients estimated by this procedure have a proportional hazards, or relative risk, interpretation. We estimated the models using a maximum likelihood method, carried out with the GENMOD procedure in the statistical program SAS.

We present three nested models predicting maternity leave adoption. First, we examine the effects of organizational and labor market variables. Second, we add variables representing maternity leave litigation and press coverage. Third, we add variables representing changes in the legal environment. In models not reported here, we added a time trend to rule out the possibility that the legal regime variables are picking up a simple secular change. They are not.

FINDINGS

We find little support for the feminization and legal sanction hypotheses and strong support for neoinstitutional hypotheses. First, employers responded to press coverage of changes in maternity leave law. Second, in

[4] The complementary log-log transformation takes a variable with the values between 0 and 1 and changes it to a continuous value that ranges from minus infinity to plus infinity (Allison 1996, p. 216).

the contentious legal environment that followed the EEOC guidelines, employers adopted maternity leave in large numbers; whereas after the Pregnancy Discrimination Act confirmed the legal standing of the EEOC position, employers were less likely to adopt maternity leave. We conclude that the contested administrative ruling caused susceptible employers to change their practices and that the congressional statute thus had no net effect. This supports our hypothesis about the paradoxical power of putatively weak administrative rulings and undermines the legal sanction hypothesis about the relative power of congressional statutes. Another test of the legal sanction hypothesis also fails: maternity leave litigation does not show a significant effect on employer policies. Our third finding about the legal environment is that California's affirmative action law was more effective than the federal equal protection law, as expected. In fully specified models, we found no support for the feminization thesis: industry-level feminization, and change in feminization, failed to show significant effects.

Our most striking findings can be seen by comparing models 1, 2, and 3 in table 2. To test the feminization hypothesis, we examine the sex composition of the workforce at the industry level. Whereas in model 1, both feminization and change in feminization showed significant effects, in models 2 and 3, in which we control for press coverage of maternity leave laws and for the laws themselves, the feminization measures are no longer significant. These findings support cross-sectional studies showing that feminization has little effect on maternity leave (Ingram and Simons 1995; Glass and Fujimoto 1995). To be sure, female workforce participation was rising in the 1970s when maternity leave policies began to diffuse, but it appears that legal shifts better predict organizational change. Because maternity leave was covered under sex discrimination law, an industry's particular sex composition is relatively unimportant as a predictor of adoption. It appears that employers adopted maternity leave not to compete for women workers so much as to comply with the law. This can be seen, indirectly, in figure 2—few employers, even those in feminized industries offered maternity leave before the law changed. It was not workforce feminization per se that spawned maternity leave policies.

We test the legal sanction hypothesis by comparing the effects of the PDA to those of the EEOC guidelines, and as noted above, the PDA did not show a net effect. We conclude that the EEOC guidelines had led employers who were attentive to the law to change their practices. We also test the legal sanction hypothesis by looking at maternity leave litigation counts. The litigation variable does show a positive effect in the presence of controls (in a model identical to model 2 but without the press count variable), but the press count variable washes that effect out. This is not surprising, for we saw the conjunction of the peak of press coverage (fig.

481

TABLE 2

ESTIMATES OF THE ADOPTION OF MATERNITY LEAVE

| | MODEL | | | | | |
|---|---|---|---|---|---|---|
| | Organizational and Labor Market Variables (1) | | With Press and Litigation Variables (2) | | With Legal Periods (3) | |
| Log employment ..... . ...... .. | .128* | (.063) | .146* | (.064) | .146* | (.064) |
| Age ... ..... .. ...... . .. ............... | − 004* | (.002) | −.004* | (.002) | −.004* | (.002) |
| Banking . ........ ........ . ...... .. .. | 1.231** | ( 287) | 1.254** | (.287) | 1.197** | (.289) |
| Core manufacturing ... .. ...... ... | 1.200** | (.271) | 1.138** | (.269) | 1 036** | (.271) |
| Government agencies .. .. ...... .. | 620* | ( 259) | .639* | ( 258) | .622* | ( 260) |
| Unionized . ............ ......... . ....... | .196 | (.228) | .146 | (.230) | .049 | (.236) |
| Personnel office  .. ......... . ....... | .182 | (.218) | .163 | (.216) | .193 | ( 214) |
| Benefits office ....... . ......... ...... | .545* | ( 212) | .513* | ( 212) | 554* | (.218) |
| Legal office ........ ...... ........ .. ... | −.039 | (.242) | .021 | (.245) | .069 | (.248) |
| Formalization scale ... . ....... .. | .080 | (.063) | .062 | (.064) | .058 | (.064) |
| Unemployment .. ...... ...... .. ....... | .265** | (.056) | .208** | (.069) | .050 | (0.83) |
| Change in unemployment  ..... | −.085 | ( 075) | −.001 | (.079) | − 060 | (.085) |
| Feminization  .... .......... . ...... .. | .014* | ( 006) | .012 | (.006) | .009 | ( 006) |
| Change in feminization . .. ...... | 8.669* | (3.496) | 7.247 | (3 798) | 6.871 | (4.084) |
| Maternity leave: | | | | | | |
| Litigation .. ....... . ... ..... .. ... | | | −.037 | ( 019) | −.009 | ( 028) |
| Press coverage ....... ......... . .. | | | .084** | (.029) | −.006 | ( 044) |
| 1974–79  ..... .. ....... ........ .. | | | | | 1.744** | (.509) |
| 1980–85 . ......... ........ . ......... | | | | | .419 | (.630) |
| California . ............ .. ........ ....... | | | | | 227 | ( 304) |
| 1974–79 .. ......... ............ ..... | | | | | −.237 | (.392) |
| 1980–85 ...... .. ......... . ....... . . | | | | | 1 233* | (.512) |
| Constant .. ......... .. ........ ......... | −7.331** | (.604) | −7.152** | (.639) | −6.273** | (.666) |
| Deviance ... .......... .. ........ ........ | 1,201.42 | | 1,185.99 | | 1,162.09 | |
| df ..... .. .............. ...... .... . ... .. .. ..... | 4,711 | | 4,709 | | 4,704 | |

NOTE —SEs are in parentheses
* P < 05.
** P < .01

1) and the rapid expansion of corporate maternity leave policies (fig. 2) in the mid-1970s. The volume of litigation (see fig. 4) rose when the legal standing of the EEOC position was confirmed by the PDA and remained high during the early 1980s. The spread of maternity leave slowed in this period. The results suggest that employers respond not to the objective risk of being sued but to press coverage that makes them aware of new legal standards.

After evaluating the indirect effects of legal changes, via press coverage, in model 2, we evaluate the direct effects of legal changes in model 3. With other factors controlled, the annual hazard of adopting maternity

leave is about 5.8 times higher in the period 1974–79 than in the period 1955–73. This suggests that the EEOC guidelines had a strong positive effect on the adoption of maternity leave. However, with full controls in the model, the coefficient for 1980–85 is small and nonsignificant. This suggests that after the Pregnancy Discrimination Act of 1978 was passed, employers were no more likely to add maternity leave policies than they had been in the period 1955–73.

We introduce the variable California and interact it with the two latter periods to determine whether California's 1978 maternity leave law had a net effect on adoption. California (which captures the period 1955–73) and California, 1974–79, show small and nonsignificant effects, suggesting that before 1980s, California employers were no different from New Jersey and Virginia employers. However, California, 1980–85, shows a large and significant effect, which suggests that the 1978 affirmative action law had an impact. While New Jersey and Virginia employers were no more likely to adopt maternity leave in 1980–85 than they had been in 1955–73, California employers who had not already done so were about five times as likely to adopt leaves. To test whether California was different from both other states, we introduced a parallel set of period interactions for New Jersey (in a model not reported here). The results—a significant positive coefficient for California, 1980–85, and an insignificant coefficient for New Jersey, 1980–85—show that California was unique.[5]

In sum, models 2 and 3 measure shifts in the legal environment in different ways. In model 2, we observe the effects of legal changes as mediated by press coverage. In model 3, we observe the effects of legal changes directly, with measures that capture each state-by-period policy environment. We do not expect to find that all measures are significant when included together. Press coverage does not show a significant effect in model 3, and we conclude that this is because the state-by-period interactions better capture the substantial variety in the law than does the press coverage variable.

Next, we turn to control variables. As expected, large employers are more likely to create maternity leave programs and older employers are

---

[5] After 1978, the combination of the Pregnancy Disability Act and state laws on temporary disability insurance meant that medium and large employers in both California and New Jersey were required to cover pregnancy in their disability programs (California Unemployment Insurance Code 2626 [1998], New Jersey Statutes 43:21–29 [1991]). Women in these states received some pay during their pregnancy-related absences, but these laws did not require job-guaranteed leaves per se. If state disability laws had been crucial to the spread of maternity leave, we would have seen California *and* New Jersey employers behave differently than Virginia employers. New Jersey employers did not jump to add formal leave policies after the PDA, pointing to the importance of California's affirmative action law.

less likely to do so. We speculated that industries that place a premium on employment continuity—including core manufacturing (chemicals, electrical machinery, and machinery) and finance—would be more likely to adopt maternity leave. We find consistent positive effects of location in these sectors.

We predicted that government agencies and nonprofit organizations would be quick to comply with new Civil Rights laws. While nonprofits were no different from organizations in the reference industries, and hence are included in the reference category in reported models, government agencies were significantly different.

We test the industrial relations hypotheses that unions fight for fringe benefits and that feminized unions fight for maternity leave. We found that unionized employers were not more likely to adopt maternity leave, and in results not reported here, we found that feminized, unionized employers were not more likely to do so either. However, we do not consider these findings to be definitive, because the industry variables may be capturing the effects of unions. The failure of the unionization measures may also be the result of a process that has been described by Selznick (1969), Kochan, Katz, and McKersie (1986), and Swenson (1992), whereby nonunion firms copy the industrial relations systems of their unionized peers.

We predicted that employers with personnel, legal, and benefits offices would be more likely to adopt maternity leave policies. The presence of a benefits office increases the annual risk by about 70%. Personnel and legal offices show no effects, apparently because the benefits profession led the charge to create maternity policies. In a search of professional journal indexes, we found that benefits journals carried 55 articles on maternity leave between 1970 and 1984—more than personnel, labor law, and executive management journals combined. The formalization scale shows positive, but nonsignificant, effects in each model. It appears that having a professional within the organization to advocate maternity leave, in the person of a benefits manager, is more important than having formalized personnel functions.

Economists expect that when unemployment is low in an industry, employers will offer benefits, such as maternity leave, to attract new workers and bind incumbents to the firm. Models 1 and 2 suggest, on the contrary, that when unemployment is *high*, employers adopt maternity leave. However, model 3, with legal periods controlled, suggests that this relationship is spurious. The oil crisis of the mid-1970s produced high unemployment, and this happened to coincide with the implementation of the EEOC ruling. When we control for this period, the unexpected positive effect of unemployment disappears.

CONCLUSION

Why did so many American employers create maternity leave programs in the 1970s and 1980s, well before passage of the Family and Medical Leave Act of 1993? Politicians, business leaders, and scholars have pointed to the growing presence of women in the workforce. We argue that sex discrimination law was the driving force. As federal agencies and the courts elaborated the definition of sex discrimination under the Civil Rights Act, employers added maternity leave policies to move into compliance.

We build on an insight from neoinstitutional theory—that structural weaknesses do not prevent the U.S. state from shaping employer practices. Previous studies have explored the effects of Congress's use of ambiguous language in employment law. The ambiguity of the law can lead employers, uncertain what exactly is required of them, to create elaborate compliance measures. Legal ambiguity can thus lead to strong corporate response. We have examined the consequences of another seeming weakness of the federal state, the separation of powers. The constitutional separation of powers allows all three branches of government to make law, but it also invites private parties to contest the law. To assess the effects of this legal openness on employer compliance, we compare the most vulnerable kind of law with the least vulnerable kind of law. Administrative law is weakest because it can be challenged by case law, which can, in turn, be overruled by legislation. In consequence, congressional acts appear to be more stable and powerful than administrative rulings. History has provided us with an opportunity to test this idea against a competing idea—that administrative rulings can be effective precisely because their vulnerability invites legal challenges, which can serve to promulgate them.

Perhaps our most surprising finding is that an embattled administrative ruling of uncertain legal standing caused large numbers of employers to adopt maternity leave policies. The EEOC's 1972 Guidelines on Discrimination because of Sex were challenged by employers who argued that the agency had no authority to extend sex discrimination law to cover pregnancy. Employers won this battle in the Supreme Court, yet, in the meantime, press coverage of their challenge helped to make other employers aware of the law. By the time Congress reversed the Supreme Court and codified the EEOC position in the Pregnancy Discrimination Act of 1978, those employers who were susceptible to legal pressure had already changed their policies. In brief, legal challenges to an administrative ruling led employers to comply with the ruling. We would expect to find the same process at work in other realms of business regulation, because

administrators and judges interpret the law in realms as diverse as antitrust and occupational safety, and because employers frequently use the courts to challenge those interpretations. Given the administrative branch's reluctance to monitor each and every workplace, organizations are likely to learn about new legal standards through press coverage of litigation. Minor administrative rulings can have large effects, then, even when they are eventually overturned by judges (see McCann 1994, 1998).

We find little evidence that employers adopted maternity leave in response to the objective risk of legal sanction. First, they were not more likely to comply with an uncontested congressional statute than with an embattled administrative guideline. Second, they were not sensitive to the number of federal maternity leave lawsuits. This does not mean that the law did not matter; it means that the law affected employer practice by creating a social standard and not simply by creating a calculable risk. The behaviorist assumptions underlying the law-and-economics approach suggest that agents make objective evaluations of the risk of sanction by observing what the state is doing. The sociological assumptions underlying both the neoinstitutional (Meyer and Rowan 1977; Meyer 1994; Scott 1995) and the network (White 1981; Granovetter 1985) approaches suggest that agents monitor the behavior of their peers. In this case, our findings suggest, their lookouts were benefits professionals who tracked emerging norms in the organizational field. Employers behaved as social actors, copying the behavior of others, rather than as rational calculators, making objective assessments of the risks associated with noncompliance.

In addition to examining the efficacy of administrative law, we compare the two kinds of Civil Rights law that govern employment: equal opportunity law and affirmative action law. We make a critical comparison between the federal Pregnancy Discrimination Act of 1978 and California's Fair Housing and Employment law, passed in the same year. The federal law requires employers who offer medical disability leaves to cover maternity; the California law requires employers to allow maternity leaves even if they do not offer medical disability leaves. We compare employers in New Jersey and Virginia, subject only to the federal equal opportunity law, to those in California, subject to the affirmative action law. California's affirmative action law has a significant net effect, observable in both the over-time plot and the event-history analysis. We expect the finding that affirmative action laws are more effective in changing employer practices to hold up in other settings, and indeed early labor market studies support this idea (Leonard 1984).

Some argue that affirmative action laws are not politically viable, and others argue that they can do more harm than good. In light of the effec-

tiveness of these policies, these fears are exaggerated. First, while political backlash has led some states to curtail affirmative action in higher education, common affirmative action practices in employment, such as recruiting and mentoring programs for women and minorities, remain popular with the public (Steeh and Krysan 1996, p. 132; Taylor 1995). Moreover, the newest affirmative action law—the Americans with Disabilities Act of 1990, which requires that employers accommodate the disabled—also enjoys widespread support. Second, while some fear that "special treatment" laws can harm the intended beneficiaries, our findings suggest that group-specific policies, functioning in the context of equal opportunity law, can make the workplace more hospitable for women. Recent evidence shows that paid maternity and parental leaves—expensive programs benefiting primarily women workers—have not limited women's employment in Europe (Ruhm and Teague 1997, pp. 147–48).

Public policy played an important role in the spread of maternity leave programs during the 1970s and 1980s, but it left little trace in the minds of managers, politicians, and scholars. These observers may discount the role of policy because the Supreme Court overturned the key law, but the fact that the EEOC guidelines were overturned does not mean that they did not have effects. They had powerful effects. High-profile legal challenges by employers appear to have amplified those effects, rather than to have dampened them.

The case we examine is typical in many ways. Legislation is just the beginning of lawmaking in the federal system. It is typical that the crucial policy change emerged when the administrative branch sought to interpret congressional intent (Lempert and Sanders 1986; Suchman and Edelman 1996). It is typical that affected parties challenged that interpretation (Burstein and Monaghan 1986; Burstein and Pitchford 1990). It is typical that professionals lobbied executives to comply with the new law even before legal challenges had been resolved (Edelman et al. 1992; Sutton and Dobbin 1996). And it is typical that a quarter century hence, politicians and business leaders appear to have forgotten this episode (Dobbin and Sutton 1998).

Commentators argue that employers created maternity leave policies in response to a change in the labor market, namely feminization. This claim reinforces the idea that, in the United States, the market is powerful and the state is not. That idea becomes self-fulfilling when even the advocates of social change come to believe it. We view the pattern we have documented, of a public policy inducing new business practices and then being forgotten, as evidence of the remarkable rhetorical power of the market in the United States. The inclination to see social practices as driven by market mechanisms, and to believe that practices created by

other means are artificial and illegitimate appears to have blinded many to the role of the state in improving opportunities for disadvantaged groups.

REFERENCES

Allison, Paul D. 1995. *Survival Analysis Using the SAS System*. Cary, N.C.: SAS Institute.
Ashenfelter, Orley, and James Heckman. 1976. "Measuring the Effect of an Antidiscrimination Program." Pp. 46–84 in *Evaluating the Labor Market Effects of Social Programs*, edited by Orley Ashenfelter and James Blum. Princeton, N.J.: Princeton University Industrial Relations Section.
Baron, James N., Frank R. Dobbin, and P. Devereaux Jennings. 1986. "War and Peace. The Evolution of Modern Personnel Administration in U.S. Industry." *American Journal of Sociology* 92:350–83.
Baron, James N., P. Devereaux Jennings, and Frank R. Dobbin. 1988 "Mission Control? The Development of Personnel Systems in U.S. Industry." *American Sociological Review* 53:497–514.
Bender, Marilyn. 1973. "Many Companies Revising Maternity Leave Policies." *New York Times*, December 10, pp. 61–62.
Blau, Peter M., Cecilia McHugh Falbe, William McKinley, and Phelps K. Tracy. 1976. "Technology and Organization in Manufacturing." *Administrative Science Quarterly* 21:20–40.
Blau, Peter M., and Richard A. Schoenherr. 1971 *The Structure of Organizations*. New York: Basic Books.
BNA (Bureau of National Affairs). 1967. *A Current Look at: (1) The Negro and Title VII, (2) Sex and Title VII. PPF Survey No. 82*. Washington, D.C.: BNA.
———. 1975. "Paid Leave and Leave of Absence Policies." *Personnel Policies Forum, Survey No. 111*. Washington, D.C.: BNA.
———. 1986. *Work and Family: A Changing Agenda* Washington, D.C.: BNA.
Burstein, Paul, and Kathleen Monaghan. 1986. "Equal Employment Opportunity and the Mobilization of Law." *Law and Society Review* 16:355–88.
Burstein, Paul, and Susan Pitchford. 1990 "Social-Scientific and Legal Challenges to Education and Test Requirements in Employment." *Social Problems* 37:243–57.
*Business Week*. 1976. "A Pregnancy Ruling that Could Cost $1.6 Billion." November 29, p. 41.
California Government Code 12945. 1998. *Deering's California Codes Annotated*. Dayton, Ohio: Lexis Law Publishing
California Unemployment Insurance Code 2626. 1998. *Deering's California Codes Annotated*. Dayton, Ohio· Lexis Law Publishing.
Carruthers, Bruce G., and Terence C. Halliday. 1998. *Rescuing Business: The Making of Corporate Bankruptcy Law in England and the United States*. Oxford: Clarendon Press.
Citizens' Advisory Council on the Status of Women. 1971. *Women in 1970*. Washington, D.C.: Government Printing Office.
———. 1973. *Women in 1972*. Washington, D.C.: Government Printing Office.
Costain, Anne N. 1992 *Inviting Women's Rebellion: A Political Process Interpretation of the Women's Movement*. Baltimore: Johns Hopkins University Press.
DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 35:147–60.
Dobbin, Frank, Lauren Edelman, John W. Meyer, W. Richard Scott, and Ann

Swidler. 1988. "The Expansion of Due Process in Organizations." Pp. 71–100 in *Institutional Patterns and Organizations: Culture and Environment,* edited by Lynne G. Zucker. Cambridge, Mass.: Ballinger.

Dobbin, Frank, and John R. Sutton. 1998. "The Strength of a Weak State: The Employment Rights Revolution and the Rise of Human Resources Management Divisions." *American Journal of Sociology* 104:441–76.

Dobbin, Frank, John R. Sutton, John W. Meyer, and W. Richard Scott. 1993. "Equal Opportunity Law and the Construction of Internal Labor Markets." *American Journal of Sociology* 99:396–427.

Doeringer, Peter B., and Michael J. Piore. 1971 *Internal Labor Markets and Manpower Analysis.* Lexington, Mass.: Heath.

Dullea, Georgia. 1974. "Sick Leave Pay Still an Issue where Pregnancy Is Concerned." *New York Times,* April 8, p. 41.

Edelman, Lauren. 1990. "Legal Environments and Organizational Governance: The Expansion of Due Process in the American Workplace." *American Journal of Sociology* 95:1401–40.

———. 1992. "Legal Ambiguity and Symbolic Structures: Organizational Mediation of Civil Rights Law." *American Journal of Sociology* 97:1531–77.

Edelman, Lauren B , Steven E. Abraham, and Howard S. Erlanger. 1992. "Professional Construction of the Law: The Inflated Threat of Wrongful Discharge." *Law and Society Review* 26:47–84.

Edwards, Harry T. 1973. "Sex Discrimination under Title VII: Some Unresolved Issues." *Labor Law Journal* 24 (July): 411–23.

Edwards, Mark E. 1996. "Pregnancy Discrimination Litigation: Legal Erosion of Capitalist Ideology under Equal Employment Opportunity Law." *Social Forces* 75:247–69.

Edwards, Richard. 1979. *Contested Terrain: The Transformation of the Workplace in the Twentieth Century.* New York: Basic Books.

Elving, Ronald D. 1995. *Conflict and Compromise: How Congress Makes the Law.* New York: Simon & Schuster.

Equal Employment Opportunity Commission. 1966. *First Annual Report.* Washington, D.C.: Government Printing Office.

Fligstein, Neil. 1990. *The Transformation of Corporate Control.* Cambridge, Mass.: Harvard University Press.

———. 1996. "Markets as Politics: A Political-Cultural Approach to Market Institutions." *American Sociological Review* 61:656–74.

Frank, Meryl, and Robyn Lipner. 1988. "History of Maternity Leave in Europe and the United States." Pp 3–22 in *The Parental Leave Crisis,* edited by Edward F. Zigler and Meryl Frank. New Haven, Conn.: Yale University Press.

Freeman, Richard B. 1981. "The Effect of Unionism on Fringe Benefits." *Industrial and Labor Relations Review* 34:489–509.

Freeman, Richard B., and Morris M. Kleiner. 1990 "The Impact of New Unionization on Wages and Working Conditions." *Journal of Labor Economics* 8:S8–S25.

Fryburger, L. Bruce. 1975. "Maternity Leave Policies under Title VII." *Labor Law Journal* 26 (March): 163–73.

Glass, Jennifer, and Tetsushi Fujimoto. 1995. "Employer Characteristics and the Provision of Family Responsive Policies." *Work and Occupations* 22 (4): 380–411.

Goldin, Claudia. 1986. "Monitoring Costs and Occupational Segregation by Sex: A Historical Analysis." *Journal of Labor Economics* 4:1–28.

Goodstein, Jerry D. 1994. "Institutional Pressures and Strategic Responsiveness: Employer Involvement in Work-Family Issues." *Academy of Management Journal* 37 (2): 350–82.

Gordon, David M., Richard Edwards, and Michael Reich. 1982. *Segmented Work, Divided Workers.* London: Cambridge University Press.

Gorney, Cynthia. 1984. "Justice and the Price of Pregnancy; Maternity Leave: The Issue Dividing the Feminist Camp." *Washington Post,* April 5, p. D1.

Granovetter, Mark. 1985. "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology* 91:481–510.

Guthrie, Doug, and Louise Marie Roth. 1999. "The State, Courts, and Maternity Policies in U.S. Organizations: Specifying Institutional Mechanisms." *American Sociological Review* 64:41–63.

Harrison, Cynthia. 1988. *On Account of Sex: The Politics of Women's Issues, 1945–1968.* Berkeley: University of California Press.

Hyatt, James C. 1972. "Women Employees Seek Disability Coverage during Pregnancy Leave." *Wall Street Journal,* December 1, p. 1.

Ingram, Paul, and Tal Simons. 1995. "Institutional and Resource Dependence Determinants of Responsiveness to Work-Family Issues." *Academy of Management Journal* 38:1466–82.

Jacobsen, Joyce P., and Laurence M. Levin. 1995. "Effects of Intermittent Labor Force Attachment on Women's Earnings." *Monthly Labor Review* 118 (9): 14–19.

Jasinoski, Jerry J. 1991. "Weighing the Merits of Family-Leave Act . . . Federal Mandate Is Not the Answer." *St. Louis Post-Dispatch,* May 9, p. 3C.

Kalbfleisch, J. D., and R. L. Prentice. 1980. *The Statistical Analysis of Failure Time Data.* New York: John Wiley & Sons.

Kalleberg, Arne L., and Mark E. Van Buren. 1996. "Is Bigger Better? Explaining the Relationship between Organization Size and Job Rewards." *American Sociological Review* 61:47–66.

Kamerman, Sheila B., Alfred J. Kahn, and Paul Kingston. 1983. *Maternity Policies and Working Women.* New York: Columbia University Press.

Kessler-Harris, Alice. 1982. *Out to Work: A History of Wage-Earning Women in the United States.* New York: Oxford University Press.

Klerman, Jacob Alex, and Arleen Leibowitz. 1997. "Labor Supply Effects of State Maternity Leave Legislation." Pp. 65–85 in *Gender and Family Issues in the Workplace,* edited by Francine D. Blau and Ronald G. Ehrenberg. New York: Russell Sage.

Kochan, Thomas A., Harry C. Katz, and Robert B. McKersie. 1986. *The Transformation of American Industrial Relations.* New York: Basic Books.

Lempert, Richard, and Joseph Sanders. 1986. *An Invitation to Law and Social Science: Desert, Disputes, and Distribution.* New York: Longman.

Leonard, Jonathan S. 1984. "The Impact of Affirmative Action on Employment." *Journal of Labor Economics* 2:439–63.

Lincoln, James R., and Arne L. Kalleberg. 1985. "Work Organization and Workforce Commitment: A Study of Plants and Employees in the U.S. and Japan." *American Sociological Review* 50:738–60.

*Los Angeles Times.* 1972. "Rules to Bar Job Bias Based on Pregnancy." *Los Angeles Times,* April 1, sec. 3, p. 10.

McCammon, Holly. 1996. "Protection for Whom? Maximum Hours Laws and Women's Employment in the United States, 1880–1920." *Work and Occupations* 23:132–64.

McCann, Michael W. 1994. *Pay Equity Reform and the Politics of Legal Mobilization.* Chicago: University of Chicago Press.

———. 1998. "How Does Law Matter for Social Movements?" Pp. 76–108 in *How Does Law Matter?* edited by Bryant G. Garth and Austin Sarat. Evanston, Ill.: Northwestern University Press, American Bar Foundation.

Meyer, John W. 1994. "Rationalized Environments." Pp. 28–54 in *Institutional Environments and Organizations: Structural Complexity and Individualism,* edited by W. Richard Scott and John W. Meyer. Thousand Oaks, Calif.: Sage.

Meyer, John W., and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83: 340–63.

Meyer, John W., and W. Richard Scott, eds. 1983. *Organizational Environments: Ritual and Rationality.* Beverly Hills, Calif.: Sage.

Milkman, Ruth. 1987. *Gender at Work: The Dynamics of Job Segregation by Sex during World War II.* Urbana: University of Illinois Press.

Milliken, Frances J., Luis Martins, and Hal Morgan. 1998. "Explaining Organizational Responsiveness to Work-Family Issues: The Role of Human Resource Executives as Issue Interpreters." *Academy of Management Journal* 5:580–92.

*National Underwriter.* 1974. "Court Finds G.E. Guilty of Sex Bias in Denial of Disability Benefits " *National Underwriter,* April 20, pp. 1–4.

New Jersey Statutes 43:21–29. 1991. *New Jersey Statutes Annotated.* St. Paul, Minn.: West Publishing.

*New York Times.* 1972. "Job Unit Widens Women's Rights." *New York Times,* March 31, p. A1.

O'Connell, Martin. 1990. "Maternity Leave Arrangements: 1961–85." *Work and Family Patterns of American Women, Current Population Reports.* Special Studies Series P-23, no. 165.

Osterman, Paul. 1995. "Work/Family Programs and the Employment Relationship." *Administrative Science Quarterly* 40:681–700.

Pear, Robert. 1985. "Should Leaves for New Parents Be Mandatory? Mr. Klein: No Concern of Government." *New York Times,* December 29, sec. 4, p. 16.

Posner, Richard A. 1974. "Theories of Economic Regulation." *Bell Journal of Economics* 5:335–58.

———. 1997. "The Rise and Fall of Administrative Law." *Chicago-Kent Law Review* 72:953–63.

Rhode, Deborah L. 1989. *Justice and Gender: Sex Discrimination and the Law.* Cambridge, Mass.: Harvard University Press.

Roy, William G. 1997. *Socializing Capital: The Rise of the Large Industrial Corporation in America.* Princeton, N.J.: Princeton University Press.

Ruhm, Christopher J., and Jackqueline L. Teague. 1997. "Parental Leave Policies in Europe and North America." Pp. 133–56 in *Gender and Family Issues in the Workplace,* edited by Francine D. Blau and Ronald G. Ehrenberg. New York: Russell Sage.

Schatz, Ronald W. 1983. *The Electrical Workers: A History of Labor at General Electric and Westinghouse, 1923–1960.* Urbana: University of Illinois Press.

Scott, W. Richard. 1995. *Institutions and Organizations.* Thousand Oaks, Calif.: Sage.

Selznick, Philip. 1957. *Leadership in Administration.* New York: Harper & Row.

———. 1969. *Law, Society and Industrial Justice.* New York: Russell Sage.

Silverman, Charlotte. 1943. "Maternity Policies in Industry." *The Child* 8:20–24.

Skrentny, John D. 1996. *The Ironies of Affirmative Action: Politics, Culture, and Justice in America.* Chicago: University of Chicago Press.

———. In press. "Friends in High Places: The Women's Movement in Costain's Inviting Women's Rebellion " In *Opportunistic Protest?* edited by Jeff Goodwin.

Spaeth, Joe L., and Diana P. O'Rourke. 1996. "Design of the National Organizations Study." Pp. 23–44 in *Organizations in America,* edited by Arne L. Kalleberg, David Knoke, Peter V. Marsden, and Joe L. Spaeth. Thousand Oaks, Calif.: Sage.

Steeh, Charlotte, and Maria Krysan. 1996. "The Polls—Trends. Affirmative Action and the Public, 1970–1995." *Public Opinion Quarterly* 60:128–58.

Stinchcombe, Arthur. 1965. "Social Structure and Organization." Pp. 142–93 in *Handbook of Industrial Organization,* edited by James G. March. Chicago: Rand McNally.

Suchman, Mark C., and Lauren B Edelman. 1996. "The New Institutionalism in Organizational Analysis." *Law and Social Inquiry* 21:903–41.

Sunstein, Cass R. 1990. *After the Rights Revolution: Reconceiving the Regulatory State.* Cambridge, Mass.: Harvard University Press.

———. 1996. "The APA at Fifty: Health-Health Tradeoffs." *University of Chicago Law Review* 63:1533–71.

Sutton, John R., and Frank Dobbin. 1996. "The Two Faces of Governance: Responses to Legal Uncertainty in U.S. Firms, 1955 to 1985." *American Sociological Review* 61:794–811.

Sutton, John R., Frank Dobbin, John W. Meyer, and Richard Scott. 1994. "The Legalization of the Workplace." *American Journal of Sociology* 99:944–71.

Swenson, Peter. 1992. "Solidarity or Segmentation: Employer Strategies and Labor Supply in Sweden and the United States." Paper presented at the annual meeting of the Industrial Relations Research Association, New Orleans, June.

Taylor, Marylee. 1995. "White Backlash to Workplace Affirmative Action: Peril or Myth." *Social Forces* 73:1385–1414.

United States Code 42: 2000e(k) 1994. *United States Code Annotated.* St. Paul, Minn.: West Publishing.

UPI (United Press International). 1984. "Maternity-Leave Law Found Illegal by Judge." *New York Times,* March 22, p. C5.

Vogel, Lise. 1993. *Mothers on the Job: Maternity Policy in the U.S. Workplace.* New Brunswick, N.J.: Rutgers University Press.

Waldfogel, Jane. 1997. "Working Mothers Then and Now: A Cross-Cohort Analysis of the Effects of Maternity Leave on Women's Pay." Pp 92–126 in *Gender and Family Issues in the Workplace,* edited by Francine D. Blau and Ronald G. Ehrenberg. New York: Russell Sage.

*Wall Street Journal.* 1972. "EEOC Declares Illegal Employer Policies That Discriminate on the Basis of Pregnancy." *Wall Street Journal,* April 3, p. 6.

Walsh, Ethel Bent. 1974. "Sex Discrimination and the Impact of Title VII." *Labor Law Journal* 25 (March): 150–54.

*Washington Post.* 1996. "Campaign '96: Transcript of the Second Presidential Debate." *Washington Post,* October 17, p. A12.

White, Harrison C. 1981. "Where Do Markets Come From?" *American Journal of Sociology* 87:517–47.

Williamson, Oliver. 1981. "The Economics of Organization: The Transaction Cost Approach." *American Journal of Sociology* 87:548–77.

# Networks, Dynamics, and the Small-World Phenomenon[1]

Duncan J. Watts
*Santa Fe Institute*

The small-world phenomenon formalized in this article as the coincidence of high local clustering and short global separation, is shown to be a general feature of sparse, decentralized networks that are neither completely ordered nor completely random. Networks of this kind have received little attention, yet they appear to be widespread in the social and natural sciences, as is indicated here by three distinct examples. Furthermore, small admixtures of randomness to an otherwise ordered network can have a dramatic impact on its dynamical, as well as structural, properties—a feature illustrated by a simple model of disease transmission.

## INTRODUCTION

The small-world phenomenon (Milgram 1967; Pool and Kochen 1978) has long been an object of popular fascination and anecdotal report. The experience of meeting a complete stranger with whom we have apparently little in common and finding unexpectedly that we share a mutual acquaintance is one with which most of us are familiar—"It's a small world!" we say. More generally, most people have at least heard of the idea that any two individuals, selected randomly from almost anywhere on the planet, are "connected" via a chain of no more than six intermediate acquaintances, a notion made popular by the Broadway play (and later movie) *Six Degrees of Separation* (Guare 1990).

But is this phenomenon merely the confluence of unlikely coincidence and curious anecdote, or is it actually indicative of the underlying structure of modern social networks and, hence, not unlikely at all? Furthermore, if the small-world phenomenon does turn out to be a deep feature

of the social world, in what other contexts can it arise (such as telecommunications and neural networks), and what mechanism drives it? Finally, does its presence in the real world have any implications for the dynamical properties of the networks in which it occurs?

An explanation of the phenomenon, and more generally, a framework for examining the properties of networks consisting of very many components, is of general sociological interest: Many social metrics, such as status (Harary 1959; Burt 1982) and power (Coleman 1973), and social processes, such as the diffusion of innovations (Rogers 1995) and transmission of influence (Friedkin 1990), are usefully represented in terms of networks of relationships between social actors, be they individuals, organizations, or nations. Indeed, the theory of social networks is one that has seen extensive development over the past three decades, yielding multiple measures both of individual significance, such as centrality (Freeman 1979, 1982; Friedkin 1991), and of network efficiency (Yamaguchi 1994a), which may elucidate nonobvious phenomena such as "key players" in an organization or its optimal structure for, say, information diffusion. Frequently, however, this research assumes linear models of social processes,[2] such as Markov models of diffusion, and is generally applied to networks that consist of a relatively small number of components.[3] While many of the measures defined in the literature can in principle be applied to networks of arbitrary size and structure, the computational costs of doing so may be prohibitive (such as for Freeman's [1979] betweenness centrality), and the benefits are at any rate unclear if the process of interest is inherently *nonlinear*, as is the case for information (or disease) contagion models involving threshold (Granovetter 1978; Arthur and Lane 1993) or refractory (Murray 1993, chap. 19) effects. Hence, the problem of analyzing efficiently the structure of extremely large networks (in which components may easily number in the hundreds of thousands, or more), and modeling the effects of structure on nonlinear dynamical processes, remains relatively unexplored. This article presents one possible approach to these very general problems by postulating a model of network formation that is sufficiently flexible to account for a wide variety of interesting cases (although not all cases, by any means, as will be pointed out) and that,

[2] "Social process" here refers to the person-to-person dynamics occurring on the network. Such *local* dynamical models are thus distinct from *global* models, such as classical homogeneous mixing models of diffusion, which are generally nonlinear (see e.g., Yamaguchi 1994b), but which do not model the underlying network explicitly.

[3] Test cases appearing in the literature typically use small numbers for ease of computation (e.g., Friedkin [1991] and Yamaguchi [1994a, 1994b] consider networks of five and seven elements respectively) and documented empirical examples rarely exceed $O(100)$ nodes.

in turn, suggests which structural parameters are the appropriate ones to study.

The problem of relating network structure to dynamics is then illustrated with a very simple model of disease spreading, which implies that the specified structural parameters are significant in determining the dynamics (usually in a highly nonlinear fashion), although probably insufficient to describe it completely. It should be pointed out that networks can affect a system's dynamical behavior in what might be termed an active and a passive sense; and that it is the passive sense that is investigated here. Active implies that the network is a device to be manipulated consciously for an actor's own ends; passive implies that the network connections themselves, in concert with blind dynamical rules, determine the global behavior of the system. The active sense has been investigated, for example, by Granovetter (1973, 1974) in the context of finding a job and Burt (1992) for maximizing social capital. The passive sense—with which this article is concerned—has been explored in systems as varied as biological oscillators (Strogatz and Stewart 1993), neural networks (Crick and Koch 1998), genetic-control networks (Kauffman 1969), epidemiology (Hess 1996*a*, 1996*b;* Longini 1988; Kretzschmar and Morris 1996), and game theory (Nowak and May 1993; Herz 1994; Cohen, Riolo, and Axelrod 1999). Before addressing any of these questions, however, it is necessary to agree on precisely what is meant by the small-world phenomenon, what is already known about it, and why such a thing should be surprising in the first place.

## PROPERTIES OF RANDOM-CLUSTERED NETWORKS

### What Is the Small-World Phenomenon?

What do we mean when we say the world is "small"? In general, there is no precise answer, but in this article, "small" means that almost every element of the network is somehow "close" to almost every other element, even those that are perceived as likely to be far away.[4] This disjuncture between reality and perception is what makes the small-world phenomenon surprising to us. But why should we perceive the world to be anything other than small in the first place? The answer to this is fourfold:

1. The network is numerically large in the sense that the world contains $n \gg 1$ people. In the real world, $n$ is on the order of billions.
2. The network is sparse in the sense that each person is connected to an average of only $k$ other people, which is, at most, on the order of thou-

---

[4] Precise definitions of "small" and "far away" require some additional terminology that is developed below.

      sands (Kochen 1989)—hundreds of thousands of times smaller than the population of the planet.

3. The network is decentralized in that there is no dominant central vertex to which most other vertices are directly connected. This implies a stronger condition than sparseness: not only must the average degree $k$ be much less than $n$, but the *maximal* degree $k_{max}$ over all vertices must also be much less than $n$.

4. The network is highly clustered, in that most friendship circles are strongly overlapping. That is, we expect that many of our friends are friends also of each other.

All four criteria are necessary for the small-world phenomenon to be remarkable. If the world did not contain many people, then it would not be surprising if they were all closely associated (as in a small town). If most people knew most other people then, once again, it would not be surprising to find that two strangers had an acquaintance in common. If the network were highly centralized—say a star—then an obvious short path would exist through the center of the star between all pairs of vertices. Finally, if the network were not clustered—that is, if each person chose their friends independently of any of their friends' choices—then it follows from random-graph theory (Bollobás 1985) that most people would be only a few degrees of separation apart even for very large $n$.[5]

But are these criteria satisfied by the real world? Given that the population of the planet is several billion and that even the most generous estimates of how many acquaintances an average person can have (Kochen 1989) is only a few thousand, then the first two criteria are likely to be satisfied. The last two conditions are harder to be sure of and certainly harder to measure, but they also seem quite reasonable in the light of everyday experience. Some people are clearly more significant players than others, but even the most gregarious individuals are constrained by time and energy to know only a tiny fraction of the entire population. What significance these individuals have must be due to other more subtle and interesting reasons. Finally, while it might be difficult to determine in practice how many of a given person's friends are also friends with each other, and even more difficult to measure this for a large population, common sense tells us that whatever this fraction is, it is much larger than that which we would expect for a randomly connected network.[6]

The first evidence that the world might indeed be small was presented over 30 years ago by the psychologist Stanley Milgram (1967). Milgram

---

[5] In fact, a random graph is a close approximation to the *smallest possible* graph for any given $n$ and $k$ (where $k_{max} \ll n$ and the variance in $k$ is not too large).

[6] To be more explicit, if the world *were* randomly connected, then one's acquaintances would be just as likely to come from a different country, occupation, and socioeconomic class as one's own. Clearly this is not the case in real life.

initiated a number of chain letters with sources in Kansas and Nebraska, to be sent to one of two targets in Boston. Each source was given the name of the target and some demographic information about them but was instructed that they could only send the letter to someone they knew by first name. If they did not know the target directly (a remote possibility), the idea was to send it to whichever of their friends they considered was most likely to. This procedure was then to be repeated, generating a chain of recipients that either reached the target or else petered out due to apathy. Of the chains that did complete, Milgram found that the median number of links in the chain was about six, thus giving rise to the famous phrase, "six degrees of separation." A later work of Milgram's (Korte and Milgram 1970) found similar results for senders and recipients in different racial subgroups, thus bolstering the claim that the world was not just small within particular socioeconomic categories but was, perhaps, small universally.

Although the first theoretical examination of the small-world phenomenon, by Pool and Kochen (1978), did not appear in published form until well after Milgram's experiment, the ideas had been in circulation for some 10 years beforehand. Pool and Kochen posed the problem in terms of the probability $(p_i)$ that two randomly selected elements of a network would be connected via a shortest path consisting of $i$ intermediaries. They calculated expected values of $p_i$ under a variety of assumptions about local network structure and stratification. They concluded, as had Milgram, that the world was probably a small one, in the sense that randomly selected pairs could generally be connected by chains of only a few intermediaries. However, their assumptions concerning network structure and the independence of connections were so restrictive that they declined to place much weight on their hypothesis. Little progress has been made on this work since, and Pool and Kochen's conclusions remain essentially unchanged (Kochen 1989).

Another starting point for theoretical investigation of the small-world phenomenon was the study of random-biased nets, developed in the 1950s and 1960s by Anotol Rapoport and his colleagues at the University of Chicago. Motivated by the desire to understand the spread of infectious diseases, Solomonoff and Rapoport (1951) calculated the expected fraction of a randomly mixed population to be infected by a small initially infected seed. Rapoport then determined the corresponding fractions to be infected in populations where network connections exhibited increasing levels of local redundancy due to effects such as homophily, symmetry of edges, and triad-closure bias (Rapoport 1953a, 1953b, 1957, 1963; Foster, Rapoport, and Orwant 1963). More sophisticated approximations were developed subsequently by Fararo and Sunshine (1964) and, later, Skvoretz (1985) to account for differentiation of ties as well as vertices. However,

as was the case with Pool and Kochen, all these approximations focused on altering the *local structure* of the network and so showed only that significant changes in global structure will result from correspondingly significant changes in local structure. This statement is not contradicted here. Rather, what is new is that equally significant changes in global structure can result from changes in local structure that are so minute as to be effectively undetectable at the local level. This is an important distinction, as it is at the local level—and only at the local level—that individuals in a network make measurements.

## Formalization of the Small-World Phenomenon

In order to make the requisite notions precise, some definitions are borrowed from graph theory. For simplicity, the networks considered here will be represented as connected graphs, consisting solely of undifferentiated vertices and unweighted, undirected edges.[7] All graphs must also satisfy the sparseness conditions specified above.

The first statistic of interest, for a given graph, is the *characteristic path length* (L), defined here as the average number of edges that must be traversed in the shortest path between any two pairs of vertices in the graph.[8] In terms of Milgram's experiment, L would be the chain length averaged over all possible sources in the network *and* all possible targets. L then is a measure of the global structure of the graph (because, in general, determining the shortest path length between any two vertices requires information about the entire graph). By contrast, the *clustering coefficient* (C) is a measure of the local graph structure. Specifically, if a vertex $v$ has $k_v$ immediate neighbors, then this neighborhood defines a subgraph in which at most $k_v(k_v - 1)/2$ edges can exist (if the neighborhood is fully connected). $C_v$ is then the fraction of this maximum that is realized in $v$'s actual neighborhood, and C is this fraction averaged over all vertices in the graph.[9] Equivalently, C can be regarded as the probabil-

---

[7] These assumptions are unrealistic in general, as many networks of interest in both the social and natural sciences are composed of weighted and directed relationships. However, generalizations of the resulting graph statistics to account for these added complexities—although straightforward in principle—may depend on the particular application at hand. Therefore, for the purpose of constructing a broadly relevant framework, undirected, unweighted graphs are the natural starting point.

[8] Variants of L have appeared in other contexts as diverse as the status of individuals in an organization (Harary 1959), the floor plans of buildings (March and Steadman 1971), the efficiency of communications networks (Chung 1986), and even the properties of chemical compounds (Wiener 1947; Rouvray 1986).

[9] Local clustering, or variants thereof, has also appeared in the literature as a measure of network structure, originally in Davis (1967)

ity that two vertices $(u, v)$ will be connected, given that each is also connected to a "mutual friend" $(w)$.

All the graphs considered in this article will be characterized in terms of these two statistics. But in order to contextualize the results—to decide, in effect, what is "small" and what is "large," what counts as "clustered" and what does not—it is necessary to determine the *ranges* over which $L$ and $C$ can vary. Three constraints are imposed upon this exercise:

1. The population size $(n)$ is fixed.
2. The average degree $k$ of the vertices is also fixed such that the graph is sparse $(k \ll n)$ but sufficiently dense to have a wide range of possible structures $(k \gg 1)$.[10]
3. The graph must be *connected* in the sense that any vertex can be reached from any other vertex by traversing a finite number of edges.

Fixing $n$ and $k$ enables valid comparisons to be made between many different graph structures. Clearly, the largest value that $C$ can attain for *any* connected graph is $C = 1$, for a *complete graph* $(k = n - 1)$. Conversely, the minimum conceivable value of $C$ is $C = 0$ for an *empty graph* $(k = 0)$. These two graphs also have extremal length properties. This, however, is not a very instructive comparison, as it is obvious that clustering and length will change as more and more edges are added to any graph. A more interesting question is how these statistics can change simply by rearranging a fixed number of edges among a fixed number of vertices. The sparseness conditions focus our attention on the most interesting terrain from the perspective of a wide range of applications in both the social and natural sciences. That is, the network is sufficiently well connected to admit rich structure, yet each element is confined to operate within a local environment that encompasses only a tiny fraction of the entire system. Finally, by insisting that all graphs be connected, $L$ is guaranteed to be a truly global statistic. Hence, comparisons of characteristic path length are valid comparisons of global structure. Bearing in mind these conditions, the following questions present themselves:

1. What is the most clustered graph possible, and what is its characteristic path length?

---

[10] Because, in the models considered here, fluctuations in vertex degree $(k_v)$ are roughly normally distributed around $k$, this condition becomes in practice the more strict condition mentioned earlier: $(k_{max} \ll n)$. This is a qualitatively stronger constraint than $k \ll n$ because it precludes not only densely connected graphs, but also star-like graphs, which have $L \approx 2$, regardless of $n$, by virtue of one (or a few) highly centralized vertices. That star graphs are ignored in this analysis is not to imply that they are uninteresting, simply that our concern here is with *decentralized* graphs—a reasonable restriction if one considers large enough $n$ (where no one member could possibly know all others).

2. What graph has the lowest possible characteristic path length, and what is its clustering coefficient?
3. What do these results imply about the relationship between the clustering coefficient and characteristic path length of a sparse graph?

Turning first to clustering, a significant insight is that, although a connected graph can only attain the maximal value of $C = 1$ when $k = (n - 1)$, even a very sparse graph may have a clustering coefficient that is, in practice, indistinguishable from the complete case. The most clustered, sparse graph possible is what might be termed the *caveman graph,* which consists of $n/(k + 1)$ isolated cliques or "caves": that is, clusters of $(k + 1)$ vertices within which all vertices are connected to all others but between which no edges exist at all. It is easy to see that this graph has $C = 1$, on a par with a complete graph. However, it fails another required condition—that all graphs must be *connected.* Fortunately, global connectivity is an easy property to achieve in this case, simply by extracting one edge from each clique and using it to connect to a neighboring clique such that all cliques eventually form a single, unbroken loop. This *connected caveman graph* (fig. 1) can be shown to have a clustering coefficient of

$$C_{caveman} \approx 1 - \frac{6}{(k^2 - 1)},\tag{1}$$

which approaches 1 as $k$ becomes sufficiently large (without violating $k \ll n$; a detailed derivation of equation 1 is given in Watts [1999, chap. 4]).[11]

One can also calculate the corresponding characteristic path length for large $n$ and $k$ (see Watts 1999, chap. 4):

$$L_{caveman} \approx \frac{n}{2(k + 1)}.\tag{2}$$

Note that, for $n \gg k$, $L$ must necessarily be large and also increases linearly with increasing $n$. Hence, the connected caveman graph can be used as a benchmark for a "large, highly clustered graph."[12]

---

[11] The connected caveman graph does not, in fact, have the highest possible clustering coefficient for fixed $n$ and $k$. For instance, the last edge required to complete the ring in figure 1 is not required for connectivity and so can remain in its clique, thus marginally increasing $C$. Other even more clustered constructions may also be possible. Nevertheless, no graph can be constructed whose clustering exceeds that of the connected caveman graph by more than $O(1/k^2)$, which becomes vanishingly small as $k$ increases.

[12] Of course, sparse, connected networks with larger $L$ can be constructed (trivially, by severing one of the between-cluster edges to form a line of clusters instead of a ring). Such changes, however, do not affect the essential structural properties of the network—that is, its linear scaling properties—and so are not of concern here. Other, more elaborate constructions (such as a large, dense cluster trailed by a long line of vertices) are ruled out by the same regularity requirement that excluded star-like graphs from consideration. Hence, within the bounds of the model, the connected caveman graph is a plausible (albeit approximate) upper limit for $L$.
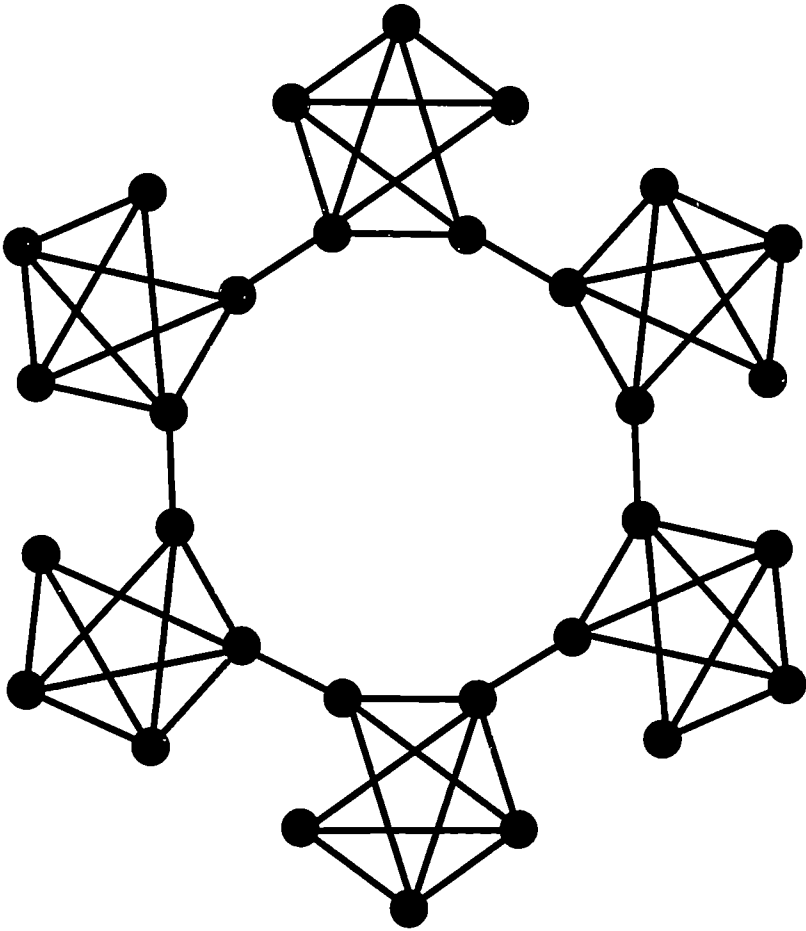
FIG. 1.—Schematic of the connected caveman construction

At the other extreme, no general, realizable structure can be shown to exhibit minimal characteristic path length for arbitrary $n$ and $k$ (Cerf et al. 1974; Bollobás 1985), but a good approximation to the theoretical lower bound is realized by a random graph (Bollobás 1985), where $kn/2$ out of all possible $n(n-1)/2$, edges are chosen at random and with equal probability. Precise formulas do not exist for $L$ and $C$ of a random graph, but in the limits of large $n$ and $k$, the corresponding asymptotic approximations are[13]

[13] For an argument justifying the approximation for $L$, see Bollobás (1985) and Bollobás and Chung (1988). The asymptotic expression for $C_{random}$ can easily be derived by

$$L_{random} \sim \frac{\ln(n)}{\ln(k)},\tag{3}$$

and

$$C_{random} \sim \frac{k}{n}.\tag{4}$$

Note that not only is $L_{random} \ll L_{caveman}$ for any $n \gg 1$, but that the scaling of $L_{random}$ is logarithmic with respect to $n$ instead of linear. Hence, as $n$ becomes larger, the discrepancy between the two extremes in length becomes increasingly pronounced (linearly in $n$). Note also that the sparse-graph condition ($k \ll n$) implies that $C_{random}$ is very small. Hence (recalling the probabilistic interpretation of the clustering coefficient), $C$ can be thought of as a simple measure of *order* in a graph—graphs with $C \gg k/n$ (like the connected caveman graph) are considered *locally ordered* (in the sense that vertices with at least one mutually adjacent vertex are likely to be themselves adjacent), and random graphs are, naturally, disordered.

The intuition that one might draw from these results is that highly clustered or locally ordered graphs necessarily have long characteristic path lengths, and conversely, graphs with short characteristic path lengths have clustering that is vanishingly small in the limit of large $n$. This is a reasonable intuition but is at odds with the (so far anecdotal) claim that the world can be small and still be highly clustered. In the absence of definitive data for the whole world, an alternative test of the small-world problem is to determine the minimum conditions that are both necessary and sufficient for the world to be small. The approach adopted here is to introduce a family of graphs that interpolates approximately between the two extremes discussed above and then to examine the intermediate region for evidence of small-world effects.

A Theory of Length Contraction in Sparse Networks

For this purpose, it is natural to consider a model that captures, in some abstract sense, the formation of social connections. A number of social network theorists have utilized the concept of a "social space" in which people exist as points separated by distances that can be measured according to some appropriately defined metric (see, e.g., Davidson 1983). Unfortunately, this approach often runs into treacherous waters due to

---

considering that a neighbor of any given vertex ($v$) has an expected probability ($k_v - 1$)/$n$ of being adjacent to another vertex ($u$) in $v$'s neighborhood.

the inherent difficulty both of characterizing the space (which is all but unknown) and defining the metric (equally so). The following three assumptions avoid these difficulties:

1. All networks can be represented solely in terms of the connections between their elements, assuming that whatever combination of factors makes people more or less likely to associate with each other is accounted for by the distribution of those associations that actually form.
2. All connections are symmetric and of equal significance. That is, a definition of what is required in order to "know" someone is defined such that either two people know each other or they do not.[14]
3. The likelihood of a new connection being created is determined, to some variable extent, by the already existing pattern of connections.

Exactly how existing connections determine new ones is a big part of the mystery. One might imagine a world in which people only become acquainted through introduction by one or more mutual friends. It is easy to see that a mechanism such as this leads inevitably to a locally ordered world (in the sense of $C \gg k/n$), the extreme case being the caveman world. At the other extreme, one might also imagine a world in which new friendships are made autonomously and at random, without regard for current friendships.[15] The end product of this tie formation process is naturally a random graph. Of course, the real world lies somewhere between these two extremes but precisely where is anybody's guess. Hence, rather than assuming some specific functional relationship between current and future friendships, let us examine a whole universe of possible "worlds" that lie between the ordered and random extremes. One way to do this in a precise and explicit fashion is through a graph construction algorithm that embodies the following features (also shown graphically in fig. 2):

1. At the ordered extreme, the propensity of two unrelated people (meaning they share no mutual friends) to be connected is very small. Once they have just one friend in common, however, their propensity to be acquainted immediately becomes very high and stays that way regardless of how many additional mutual friends they may have. In worlds like this one, it is almost a certainty that the only people anyone will ever connect to are those with whom they share at least one mutual friend. So, plotting "propensity to become friends" against "fraction of current mutual friends," the propensity starts near zero, rises very rapidly to some relatively large number (which can be normalized to one), and then plateaus.

[14] Unlike the one-way connection that often exists between, e.g., a professor and a student, or a celebrity and a fan.

[15] We may be seeing the beginnings of such a world already in the proliferation of Internet "chat lines," where complete strangers can meet, interact, and sometimes even end up marrying.
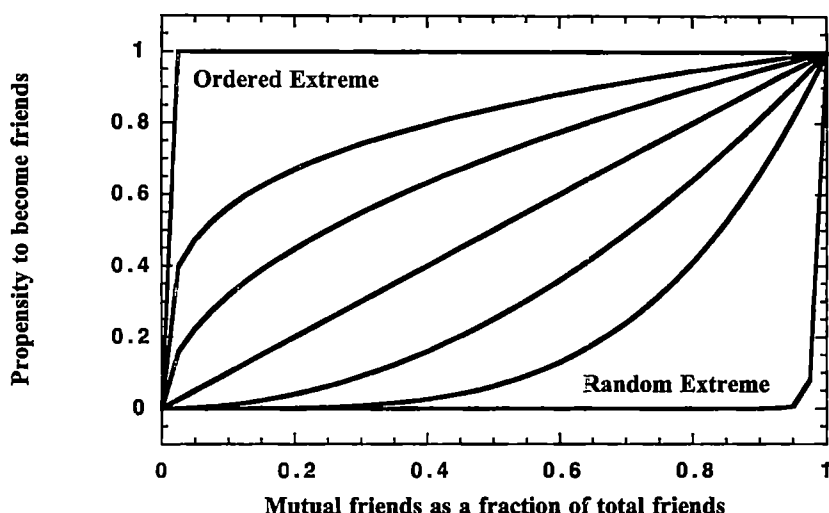
FIG. 2.—Family of functions representing the propensity of strangers to meet, given the fraction of the friends that they currently share.

2. At the random extreme, no one has much propensity to connect to anyone in particular. In this sort of world, the "propensity versus mutual friends" curve remains near zero up until the point where all friends are mutual friends.[16]

3. In between these two extremes, the propensity curve can take any one of an infinite number of intermediate forms, specified by a single, tunable parameter, where it is important only that the dependency be smooth and monotonically increasing with respect to increasing mutual friends. These conditions are satisfied by the following construction:

$$R_{i,j} = \begin{cases} 1 & m_{i,j} \geq k \\ \left[\dfrac{m_{i,j}}{k}\right]^{\alpha}(1-p)+p & k > m_{i,j} > 0, \\ p & m_{i,j} = 0 \end{cases} \tag{5}$$

where $R_{i,j}$ is a measure of vertex $i$'s propensity to connect to vertex $j$ (zero if they are already connected), $m_{i,j}$ is the number of vertices that are adjacent both to $i$ and $j$, $k$ is the average degree of the graph, $p$ is a baseline

[16] The rapid jump from near zero to near one in fig. 2 is necessitated by continuity conditions but can be rationalized in modeling terms by considering that in such a situation the two parties must have all their friends in common, in which case it is reasonable to assume that they cannot avoid meeting.

propensity for an edge $(i, j)$ to exist (set at $p = 10^{-10}$),[17] and $\alpha$ is a tunable parameter, $0 \leq \alpha \leq \infty$.[18]

*Numerical simulation.*—Equation (5) is, in principle, an abstract representation of a graph or rather, for each value of $\alpha$, an enormous (but finite) number of *potential* graphs that share certain statistical characteristics. However, it seems unlikely that the properties of these graphs, and how those properties vary with $\alpha$, can be derived in any precise analytical sense. This is an important point, because a great deal of work in graph theory is analytical. However, analytical approaches are generally confined to cases where either $n$ is small and the rules for constructing the graph are strictly deterministic or $n$ is so large that it can be treated as infinite and the rules are strictly random. Both of these extremes exhibit certain properties that simplify the situation, thus enabling analytical descriptions. The case presented by equation (5), however, falls squarely into the messy no-man's-land between these two extremes: $n$ is large but not infinite, and the rule for constructing edges is partly deterministic and partly random, where even the balance between determinism and randomness varies as $\alpha$ changes. The only manner in which such a model can be analyzed is through a rigorous process of computer-based, numerical simulation.[19] Adopting this approach, equation (5) now forms the basis for a construction algorithm,[20] which builds a graph of specified $n$, $k$, and $\alpha$.[21]

A problem that immediately rises with this "$\alpha$-model" is that, for small $\alpha$, the resulting graphs tend overwhelmingly to consist of small, isolated, and densely internally connected components. This results from a *start-up problem*—initially no edges exist in the graph, so $p$ dominates $R_{i,j}$ in equation (5), and edges form randomly until, by chance, two edges share a vertex in common. At this point, because $\alpha$ is small, the two vertices that share a mutual "friend" will almost certainly become connected at the expense of expanding their friendship networks into new territory. The fraction of pairs of vertices that are members of the same connected component grows only linearly with $k$, so connected graphs cannot be

[17] The actual value of $p$ is not important so long as it is small enough that no random edges can be expected for $\alpha = 0$ (i.e., $p \ll 2/n[n - 1]$). More specifically, the numerical results with respect to $\alpha$ do change as $p$ changes, but this dependency disappears when the model is recast in terms of the model-independent parameter $\phi$, introduced below. That is, $\phi(\alpha)$ varies with $p$, but $L(\phi)$ and $C(\phi)$ do not.

[18] Note that $\alpha$ has no physical or social significance—it is simply a parameter that enables the model to generate graphs ranging from highly ordered to highly random.

[19] All work described was conducted on either a SUN\ Sparc 20 or a DEC\ alpha 500 running $C$ under a UNIX operating system.

[20] The technical details of the algorithm are provided in the appendix.

[21] Unless otherwise stated, the parameters used to generate the results presented here are: $n = 1,000$, $k = 10$.

generated for small $\alpha$ without violating the sparse graph condition. Disconnected graphs pose a problem because they necessarily have $L = \infty$, and this makes them hard to compare with connected graphs or even each other. One way to resolve this dilemma is to build in a connected substrate before commencing the algorithm, thus ensuring that all subsequently constructed graphs will be connected. A potentially serious objection to such a step is that the properties of the resulting graphs may be so dominated by the presence of the substrate that any conclusions drawn from the model will fail to be sufficiently general to be of interest. The following constraints on the choice of substrate minimize (although do not remove) this concern:

1. It must exhibit *minimal structure,* in that no vertices are to be identified as special. This eliminates structures like stars, trees, and chains, which have centers, roots. and end-points, respectively.
2. It must be *minimally connected.* That is, it must contain no more edges than necessary to connect the graph in a manner consistent with condition 1.

The only structure that satisfies both these criteria is a topological ring. One advantage of this choice of substrate is that for sufficiently small $\alpha$, it results in graphs that resemble the connected caveman limit described above; that is, densely intraconnected clusters strung loosely together in a ring (with its attendant linear length-scaling properties). It is less obvious that the random limit also can be attained with this additional structure built in. Nevertheless, numerical evidence suggests that this is precisely the case. Hence, the ring substrate not only ensures connectivity, but also allows the model to interpolate between roughly the desired limits. More important, we will see that the results generated by the corresponding model exhibit sufficiently generic features in the intermediate regime that quite general conditions can be specified under which small-world networks should arise.

The clearest way to see this is to measure $L$ and $C$ for the $\alpha$-model, for fixed $n$ and $k$, over a range of $0 \leq \alpha \leq 20$. The following functional similarities between $L(\alpha)$ and $C(\alpha)$ are revealed (figs. 3 and 4):[22]

1. For large $\alpha$, both statistics approach their expected random-graph values.
2. At $\alpha = 0$, both $L$ and $C$ are high relative to their random-graph limits and increase (as $\alpha$ increases) to a distinct maximum at small $\alpha$.
3. Both statistics exhibit a sharp transition from their maximum values to their large-$\alpha$ limits.

----

[22] Each of the points in fig. 5 is the average value of the relevant statistic over 100 random realizations of the construction algorithm for the corresponding value of $\alpha$.
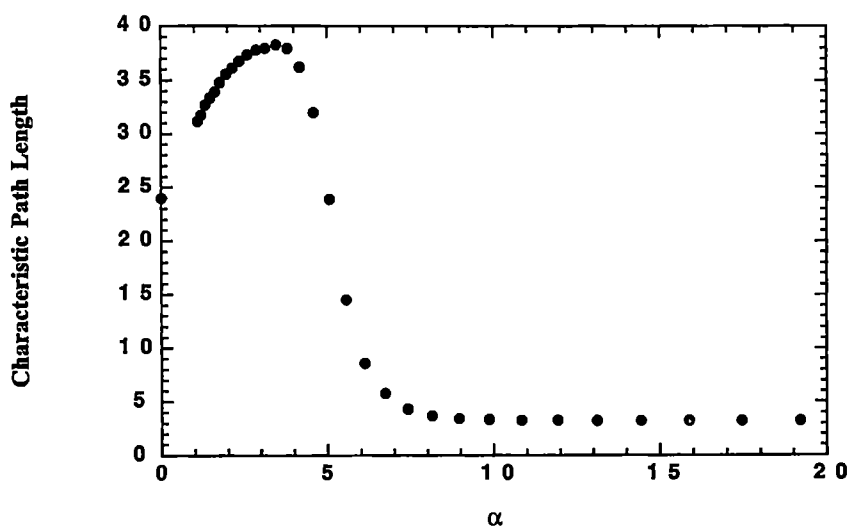
FIG. 3.—Characteristic path length ($L$) as a function of $\alpha$ for the $\alpha$-model defined by equation 5 ($n = 1,000$, $k = 10$).
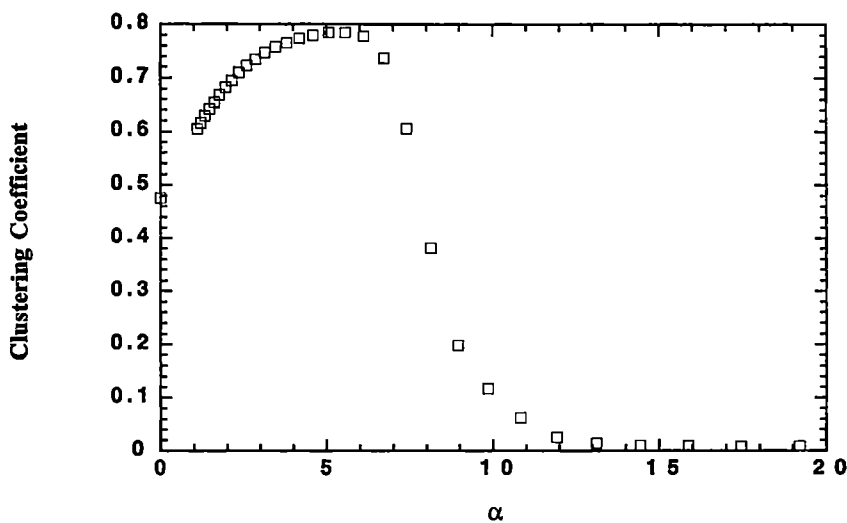


FIG. 4.—Clustering coefficient ($C$) as a function of $\alpha$ for the $\alpha$-model defined by equation 5 ($n = 1,000$, $k = 10$).
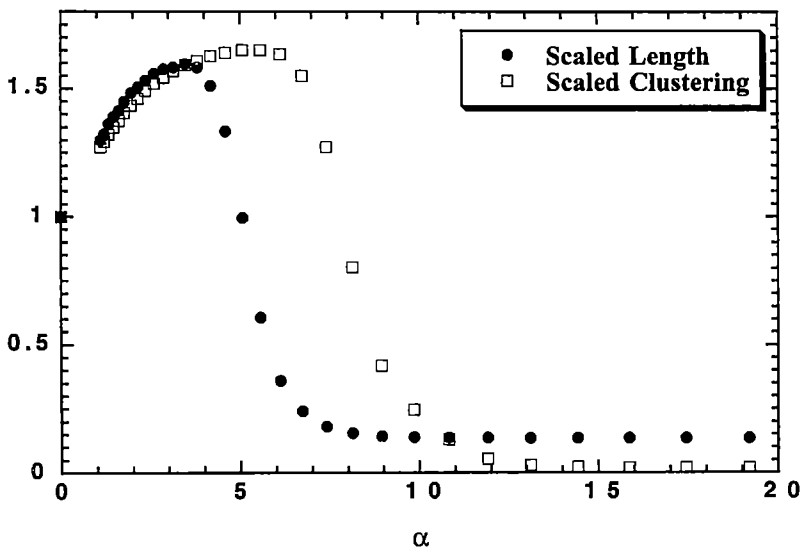
FIG. 5.—Comparison of $L(\alpha)$ and $C(\alpha)$, scaled by their corresponding values at $\alpha = 0$.

So far, the model appears to conform to the earlier statement that highly clustered graphs have large characteristic path lengths and, conversely, that graphs with short characteristic path lengths are necessarily poorly clustered.[23] However, as figure 5 indicates, $L(\alpha)$ and $C(\alpha)$ exhibit one important functional difference: the transition from large to small clustering coefficients occurs at a larger value of $\alpha$ than the equivalent length transition. The upshot of this disparity is that there exists a class of graphs in this region of $\alpha$ for which characteristic path length is small but clustering is high.[24] Thus the small-world phenomenon can be cast in graph-theoretic terms as the *coexistence* of high clustering and small characteristic path length.

DEFINITION 1.—*A small-world graph is a large*-n, *sparsely connected, decentralized graph* (n $\gg$ k$_{max}$ $\gg$ 1) *that exhibits a characteristic path*

[23] This conclusion is bolstered by the additional observation that, in the small-$\alpha$ regime, $L$ scales linearly with respect to $n$ and logarithmically with respect to $n$ for large $\alpha$, in correspondence with eqq. (2) and (3).

[24] These results remain qualitatively the same for a wide range of $n$ and $k$, strongly suggesting that they are true for *all n* and $k$, with the usual caveat $n \gg k \gg 1$. Of course, in practice, these inequalities are imprecise, the effective limits being: if $k$ is too small (in this case $k \to 2$), the substrate *will* dominate the results; and if $k$ is too large ($k \to n$), then all topologies will be equivalent.

*length close to that of an equivalent random graph* (L $\approx$ L$_{random}$), *yet with a clustering coefficient much greater* (C $\gg$ C$_{random}$).

This definition does not depend on the specifics of the graph-construction algorithm—in fact, it can be applied to any graph regardless of its construction. Nevertheless, the definition is only *interesting* if the phenomenon it describes can also be shown to be independent of the specifics of the model—in particular, the substrate. The reason for this is obvious: networks in the real world are no more likely to be constructed on ring substrates than they are to be completely ordered or completely random. This potential shortfall in the theory can be addressed in two ways. First, a variety of different substrates can be tested and their results compared with those generated above. If small-world graphs are still attainable over a significant interval of $\alpha$ values, then there is reason to think that they constitute a robust class of graphs. Second, a theoretical understanding of length contraction in partly ordered, partly random graphs may shed some light on the existence of small-world graphs and help to specify *model-independent* conditions that, if satisfied, will yield small-world graphs.

The first approach is straightforward but tedious. The same model has been tested with a number of other substrates—a two-dimensional lattice, a Cayley tree, and a random substrate. All the substrates surveyed exhibit qualitatively different properties from those of a ring, and also each other, yet $\alpha$-graphs based on all substrates invariably yield small-world graphs over an extended interval of $\alpha$.[25] The second approach—a theoretical explanation of small-world graphs—is presented in the next section.

*Shortcuts and contractions.*—Drawing intuition from the results of the $\alpha$-model, it is now possible to develop an explanation of small-world graphs in terms of a parameter that is independent of the particular model used to construct them. To motivate this approach, note that, according to equation (5) for $\alpha = 0$, newly created edges are virtually guaranteed to complete at least one triad.[26] Two vertices connected by any such edge must necessarily have been separated by a path of length two, prior to the addition of the new edge. Hence, the addition of a new edge to an $\alpha$-graph at $\alpha = 0$ contributes little in the way of length contractions, as it can connect only pairs of vertices that are already "close." In random graphs, however, this condition no longer applies, and vertices that are widely separated are as likely to become connected as those that are near neighbors. These observations lead to the following definitions:

[25] A detailed description of these substrates and the corresponding model properties is presented in Watts (1999, chap. 3).

[26] This is aside from the negligible propensity $p$ to make a random connection
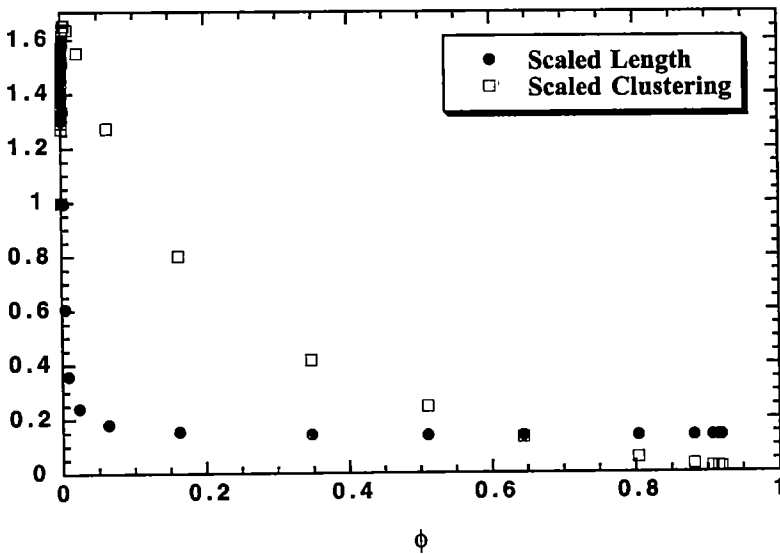
FIG. 6.—Comparison of $L(\phi)$ and $C(\phi)$, scaled by their corresponding values at $\phi_0 = \phi(\alpha = 0) = 0$.

DEFINITION 2.—*The range r of an edge is the distance between the two vertices that the edge in question connects when the edge itself has been deleted. Equivalently, range can be thought of as the second-shortest path length between two connected vertices (where the shortest path length is necessarily one).*

DEFINITION 3.—*A shortcut is any edge with a range r > 2.*

DEFINITION 4.—*The parameter $\phi$ is the fraction of all edges in the graph that are shortcuts.*

Figure 5 can now be replotted against $\phi$ instead of $\alpha$. The result, in figure 6, demonstrates not only the previous result that it is possible for highly clustered graphs to have small characteristic path lengths, but also that this happens principally for small $\phi$.[27]

[27] For very small $\phi$ (too small to resolve on the linear scale of fig. 6), there is an apparent increase in both $L$ and $C$ as $\phi$ increases—analogous to the humps in fig. 5. The basis of this effect is that, for small but nonzero $\alpha$, an increase in $\alpha$ results in higher local clustering, and so previously overlapping neighborhoods can become distinct, yielding edges with $r > 2$ that do not connect previously distant parts of the graph. This is indeed a practical problem with detecting shortcuts for this particular construction algorithm, but it occurs only at very small $\phi$—below the value at which the small-world phenomenon is relevant. Hence, it does not affect any of the results stated here.

The intuitive explanation for this additional observation is that, for small $\phi$, the characteristic path length of the graph is large. Hence, the introduction of a single shortcut is likely to connect vertices that were previously widely separated. This shortcut then contracts the distance not only between that pair of vertices, but also between their immediate neighborhoods, their neighborhoods' neighborhoods, and so on. Thus, one single shortcut can potentially have a highly nonlinear impact on $L$. By contrast, the clustering coefficient $C$ is only reduced in a single neighborhood—as the result of one less triad being formed—and so the decrease in $C$ can be at most linear in $\phi$. This nonlinear (global) versus linear (local) impact of shortcuts enables large $C$ to coexist with small $L$ at small values of $\phi$. But once $\phi$ becomes large, $L$ has already decreased to a small value, and so subsequent shortcuts can do little to reduce it further. Thus $L(\phi)$ must approach its random graph limit asymptotically, as shown in figure 6.

The parameter $\phi$ turns out to have explanatory power even beyond the specific construction presented above. Other models of graphs that interpolate between ordered and random limits also exhibit length and clustering properties that can be understood in terms of shortcuts, the key requirement being that shortcuts be permitted to connect vertices that are separated by distances on the order of the size of the entire graph. Thus, it can be conjectured that any graph with the property $n \gg k_{max} \gg 1$, which exhibits (a) a clustering coefficient $C \gg k/n$ and (b) a small fraction of long-range shortcuts, will be a small-world graph.

If this conjecture is true, then small-world graphs can be realized by a great many construction algorithms, of which the $\alpha$-model is but one. However, it also suggests that there are many kinds of partly ordered, partly random graphs in which the small-world phenomenon will not occur. The key criterion that the small fraction of introduced shortcuts be "long range" is really equivalent to the statement that new connections be determined without regard to any kind of *external length scale* imposed upon the graph (such as by explicitly disallowing connections to be made between vertices that are separated by greater than a certain physical distance). It is the independence of any external length scale that enables a tiny fraction of shortcuts to collapse the characteristic path length of the system to near its asymptotic, random-graph value, without significantly reducing the corresponding clustering coefficient. This constraint implies that the small-world phenomenon is unlikely to be exhibited by networks whose connectivity is determined solely by physical forces, which imply corresponding length scales.

Although the above conjecture appears to be a sufficient condition for the existence of small-world networks, it turns out that it is not necessary. In other words, it is possible to contract distances in large graphs with a
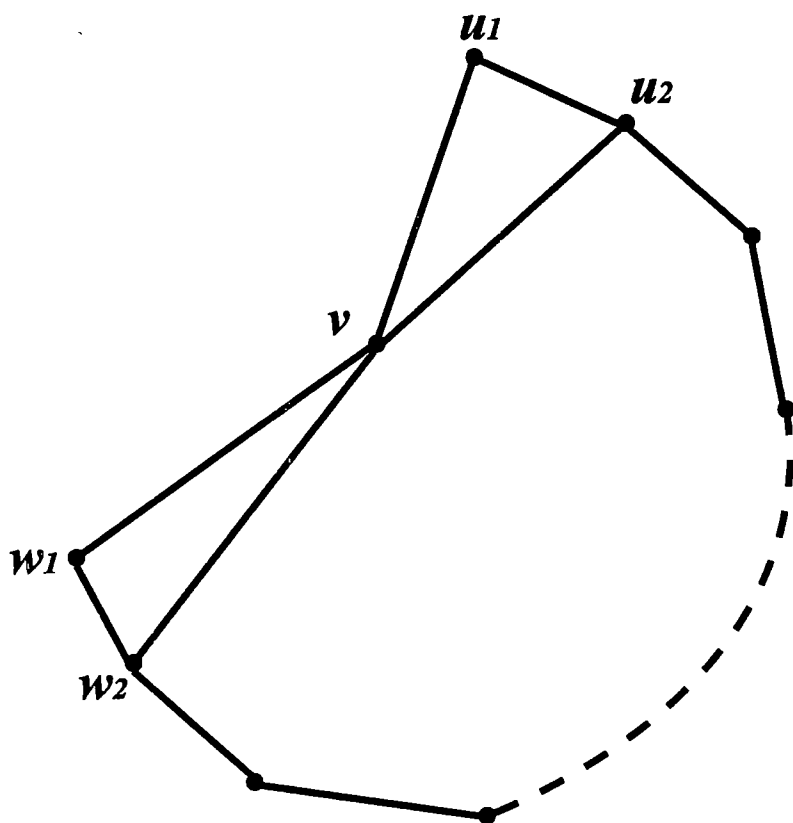
FIG. 7.—Schematic of a contraction: vertex $v$ contracts the path length between groups $u$ and $w$, but no edges are shortcuts.

negligible effect on the clustering, without using any shortcuts at all. The simplest example of such a situation is detailed in figure 7, from which it is obvious that groups of vertices are being brought closer together by virtue of a single common *member* but that none of the edges involved is a shortcut. Fortunately, a simple modification of the definition of a shortcut is adequate to capture this new scenario:

DEFINITION 5.—*A contraction occurs when the second-shortest path length between two vertices, sharing a common neighbor, is greater than two. In other words, a contraction is a pair of vertices that share one and only one common neighbor.*

DEFINITION 6.—*By extension, ψ can be defined as the fraction of pairs of vertices with common neighbors that are contractions.*

The results displayed in figure 7 can be expressed in terms of ψ (see
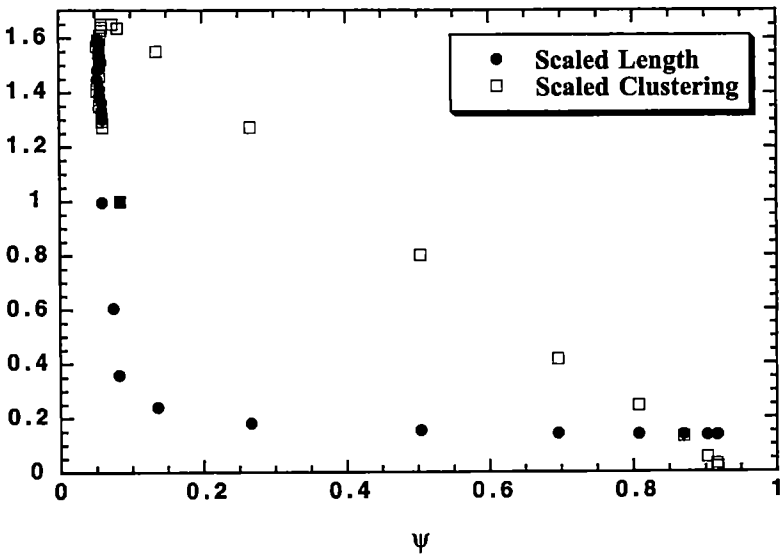
FIG. 8.—Comparison of $L(\psi)$ and $C(\psi)$, scaled by their corresponding values at $\psi_0 = \psi(\alpha = 0) \approx .08$.

fig. 8), which again shows the coexistence of small $L$ and large $C$ over a large range of $\psi$.[28] It follows from these definitions that a shortcut is simply a special case of a contraction in which one (or both) of the "groups" consists of a single vertex. Long-range contractions are thus a more general mechanism than shortcuts for generating small-world networks. In principle though, both shortcuts and contractions achieve essentially the same end—to connect what would otherwise be distant parts of a large, sparse graph with a large characteristic path length. Hence, because shortcuts are conceptually simpler and require less computational effort, all results in this article are stated in terms of $\phi$, bearing in mind that they can be recast in terms of $\psi$ if necessary.

Evidence of the Small-World Phenomenon in Real Networks

Milgram's experiment suggested that the idea of the small-world phenomenon is at least plausible for a real social network the size of the U.S. population. But the precise mapping out of social networks of this magnitude is a practical impossibility, due to the ambiguities inherent in defining

[28] Note that, unlike $\phi$, $\psi$ is never zero, even for the most clustered graphs.

both a member of the network and also what constitutes a friendship. Furthermore, tracing experiments like Milgram's face a number of technical obstacles, which he outlines (Milgram 1967); the main problem being that one can never be sure that the chain of intermediaries actually traced between two people is the shortest one possible.[29] We can now understand, in terms of shortcuts, at least one reason why this may be so. As noted in the previous section, small-world graphs occur for values of $\phi$ at which most vertices have no shortcuts at all. For example, for $n = 1,000$ and $k = 10$, $\phi = .01$ is sufficient for $L$ to be indistinguishable from that of a random graph. If only one in a hundred edges is a shortcut, however, then (for $k = 10$) about 90% of all vertices will have no shortcuts in their local neighborhood. This absence of global information at the local scale poses significant problems for effective tracing of shortest path lengths in a network, as such an exercise requires knowledge not just of one's friends, but of one's friends' friends, and so on. If shortcuts exist, but only outside of one's local network vicinity, it becomes extremely difficult to utilize them consciously and thus construct an optimal path.

For these reasons, along with the practical difficulty associated with the empirical estimation of even local parameters like $k$ and $C$ (Kochen 1989), direct resolution of the small-world phenomenon in the actual social world seems unlikely. However, there is nothing about the small-world graph definition above that demands the graph in question represent a social network. In fact, one of the useful aspects of the corresponding conjecture is that it is quite general, specifying neither the nature of the vertices and edges, nor a particular construction algorithm required to build the graph. If it is true, then many real networks, satisfying the required $n$ and $k$ conditions, should turn out to be small-world networks. Examples of large, sparse graphs are easy to think of (neural networks, large organizations, citation databases, etc.) but difficult to obtain in the required format where both vertex and edge sets are precisely defined and completely documented. Nevertheless, three scientifically interesting examples are presented below.

The first example is the collaboration graph of feature-film actors. In the *actor collaboration graph* (Tjaden 1997), a vertex is defined as a cast member of any feature film registered on the Internet movie database,[30]

---

[29] On the other hand, as White (1970) points out, longer chains tend not to complete, biasing the sample toward shorter chains and leading to a corresponding underestimate of length.

[30] The Internet Movie Database (http·//www us.imdb com) lists the cast members of all films, of all nationalities, since 1898. The graph studied in this article is actually the largest connected compenent of the entire graph, consisting of about 90% of all actors listed in the IMDB as of April 1997.

TABLE 1

CHARACTERISTIC PATH LENGTH ($L$) AND CLUSTERING
COEFFICIENT ($C$) FOR THREE REAL NETWORKS

|  | $L_{Actual}$ | $L_{Random}$ | $C_{Actual}$ | $C_{Random}$ |
|---|---|---|---|---|
| Movie actors . .. | 3.65 | 2.99 | .79 | .00027 |
| Power grid .. .. | 18 7 | 12.4 | .080 | .005 |
| C. elegans .. ...... | 2 65 | 2 25 | 28 | .05 |

and an edge represents two actors appearing in the same movie. This structure is interesting, as it is a simple case of a large ($n = 226,000$), sparse ($k = 61$) social network. It is also reminiscent of the collaboration graph of mathematicians that is traditionally centered on Paul Erdös (Grossman and Ion 1995).[31] The second example—the *western states power graph* ($n = 4,941$, $k = 2.94$)—represents the power-transmission grid of the western United States and is relevant to the efficiency and robustness of power networks (Phadke and Thorp 1988). Vertices represent generators, transformers, and substations, and edges represent high-voltage transmission lines between them. The final example is that of the neural network of the nematode *C. elegans* (White, Thompson, and Brenner 1986; Achacoso and Yamamoto 1992)—the sole example of a completely mapped neural network. For the *C. elegans graph* ($n = 282$; $k = 14$), an edge joins two neurons if they are connected by either a synapse or a gap junction. All edges are treated as undirected, and all vertices as identical, recognizing that these are crude approximations from a biological perspective.

Table 1 shows a comparison between $L$ and $C$ for each of these graphs and also $L$ and $C$ for random graphs with the same $n$ and $k$. Note that, in each case, the characteristic path length is close to that of the equivalent random graph, yet the clustering coefficient is consistently much greater. A graphical way to view the same relationship is presented in figure 9. Here, $L$ is plotted versus $C$ for the three real graphs, and also for equivalent (that is, with same $n$ and $k$) connected caveman graphs, where in each case, the statistics have been normalized by their corresponding random-graph values. From this picture, it is clear that not only are the real networks statistically distinct from both their random and caveman equivalents, but they are all distinct in the *same way*. Furthermore, their

---

[31] This collaboration graph would also be an interesting case to examine. Unfortunately, the only data available is that in the immediate neighborhood of Paul Erdös, and this is not sufficient to draw any conclusions about its global structure.
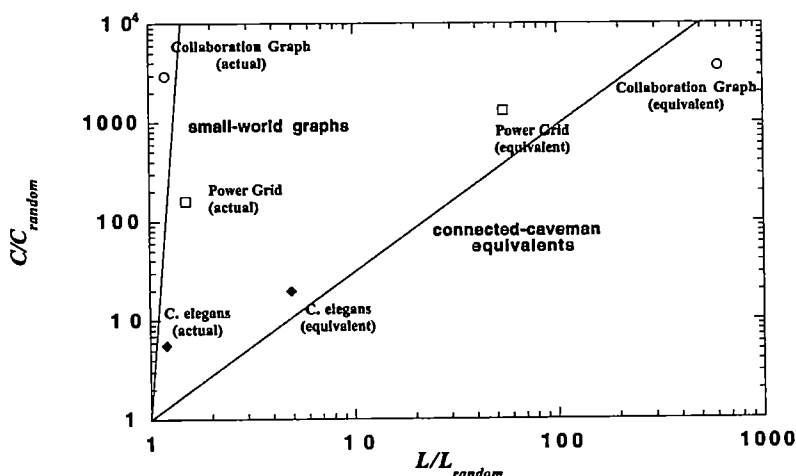
FIG. 9.—$L$ and $C$ statistics for the three real networks (the actor collaboration graph, the western states power graph, and the *C. elegans* graph) and their connected caveman equivalents (same $n$ and $k$), scaled by their values for equivalent random graphs.

combination of small $L$ and large $C$ cannot be explained as a linear interpolation between the two extremes—in other words, clustering and length do not vary in a commensurate fashion. As figure 10 shows, however, this combination of properties can be replicated by the $\alpha$-model. It is clear that, as $\alpha$ increases, the $\alpha$-statistics remain at first clustered along the (high-$L$, high-$C$) diagonal and then sharply depart from the diagonal to become small-world graphs, decreasing rapidly in characteristic path length while remaining almost constant in terms of the clustering coefficient. As $\alpha$ increases to reach the clustering transition, $C$ decreases rapidly for $L$ fixed near its asymptotic limit, until the random limit is reached at large $\alpha$. These results indicate that the small-world phenomenon is not just a property of an abstract class of hypothetical graphs, but arises in real networks. Furthermore, it is not specific to a particular kind of network or restricted to a certain size range. The three graphs examined span a range of three orders of magnitude in $n$ and represent completely different actual networks, yet all are small-world graphs in the sense defined above. Finally, not only are all three graphs small-world graphs in the broad sense of high clustering coexisting with small characteristic path length, but the relationship between $L$ and $C$ in the real graphs is consistent with the corresponding statistics of the $\alpha$-graph model.
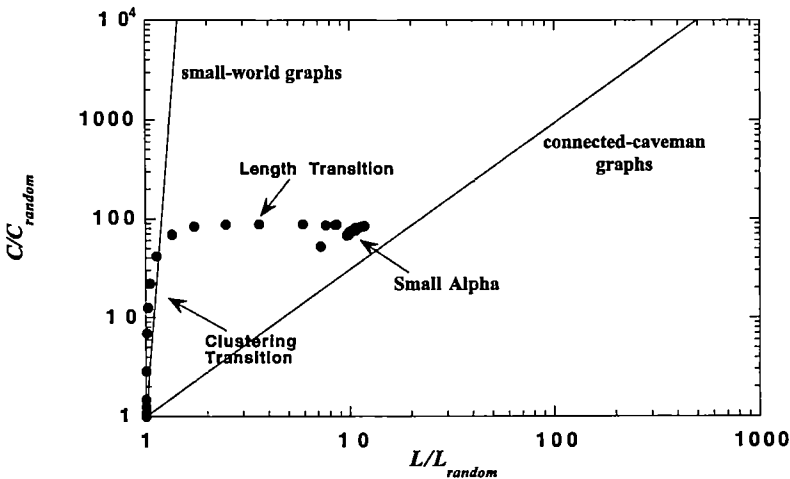
FIG. 10.—$L$ and $C$ statistics for the $\alpha$-model ($n = 1,000$, $k = 10$), scaled by their values at the random-graph limit (in practice, this is taken as $\alpha = 20$).

## DYNAMICAL SYSTEMS ON SMALL-WORLD GRAPHS

Having established that a set of relatively tiny perturbations to the local structure of a highly clustered graph can have a dramatic impact upon its global structural properties, it is natural to ask whether or not the same changes can also affect the behavior of dynamical systems that are coupled according to such a graph. This is a topic that is directly relevant to the social sciences: the role of social structure in generating globally observable, dynamical features. So far, structure has been treated as an autonomous feature of networks and defined narrowly in terms of sparse, undirected graphs. This has paid off by yielding some robust statements about the small-world properties of a general class of graphs that are partly ordered and partly random, and that seem to reflect some of the features of real networks. However, a greater issue is to understand the relationship between structure and dynamics.

As such, the following discussion builds not only on issues of network structure, but also upon a whole literature of distributed dynamical systems, in which systems are often assumed to be completely connected (the most tractable case). Where sparsely coupled systems are considered, the relevant coupling topology is usually assumed to be either completely ordered (e.g., a low-dimensional cubic lattice) or completely random. As emphasized earlier, real networks are likely to be sparse and may combine significant elements both of order and randomness, with resultant proper-

517

ties, like the small-world phenomenon, that cannot be captured by either of these approximations. Whether or not this structural oversight turns out also to be an oversight from the perspective of dynamical systems is not obvious. This article offers a more specific way to cast the issue at hand: do significant new phenomena emerge when the connectivity of distributed dynamical systems are modeled on the graphs presented earlier?

### Disease Spreading in Structured Populations

A simple case of a dynamical system is that of a disease spreading from a small seed of initiators into a much larger population whose structure is prescribed by some underlying graph.[32] The bulk of previous work on the spread of diseases focuses on populations in which uniform mixing is assumed between elements (see, e.g., Murray 1991). This is an important assumption because it enables population structure to be ignored, thus greatly simplifying the analysis. Some work, however, has grappled with the issue of population structure. Kareiva (1990) reviews a number of such attempts (which he calls "stepping-stone" models), and May and colleagues (Hassell, Comins, and May 1994; May 1995) have considered various parasite-host problems on two-dimensional grids of discrete but homogeneous patches. Both Kareiva and May conclude that the introduction of spatial structure can significantly affect both the population size and its susceptibility to parasites and disease. A similar approach has been used by Hess (1996a, 1996b) to compare virus transmission among subpopulations that are connected according to various simple topologies such as a ring and a star. Also, Sattenspiel and Simon (1988) have considered a detailed model of the spread of an infectious disease in a structured population in which different connective arrangements between the subpopulations are compared. Finally, Longini (1988) has utilized real airline network data in order to model the 1968 global outbreak of influenza. None of this work, however, goes on to treat the global dynamical properties of the system explicitly as a function of the structure. The models of Sattenspiel, Simon, and also Hess do consider different types of connectivity between subpopulations, but they consider isolated topologies, as opposed to a continuum, and their choices reflect those cases which, in the context of this article, are extremal (such as a ring versus a random graph). Extremal cases are certainly natural to consider, but if it is true that real

---

[32] Here I discuss disease spreading because of its obvious public health relevance, but qualitatively similar dynamics could describe the spread of other kinds of contagion such as ideas, rumors, fashion, or even crime (Gladwell 1996).

networks exhibit important properties of both ordered and random networks, then it is important to consider the intermediate regime as well.

*Dynamics as a function of structure.*—In this sense, the work most closely related to the approach taken here is that by Kretzschmar and Morris (1996), who analyze the spread of a disease—both in terms of extent and time scale—as a function of the overall concurrency of relationships in the population. Roughly speaking, they examine a family of graphs that interpolates between a world of exclusively monogamous (but randomly formed) relationships and an unconstrained random graph, in which concurrent relationships are likely to form. They determine that increased concurrency of relationships significantly increases the extent of the disease and its rate of spread, even when the total number of relationships in the population is held constant. Essentially, this result stems from the increasing connectedness of extremely sparse ($k = 1$) graphs: serial monogamy yields almost completely disconnected graphs (in the sense that no connected components larger than dyads can exist), but random graphs with the same number of edges exhibit relatively large connected components. Hence, as Kretzschmar and Morris conclude, it is the size of the largest connected component that drives the spread of disease across the population. The approach here also examines the effects on disease spread of changing the distribution of a fixed number of edges over a fixed number of vertices. It is different, however, in that Morris and Kretschmar consider changing concurrency in an otherwise randomly mixing population with $k = 1$, while here the amount of randomness is varied in a connected population with $k \gg 1$. Hence, any observed differences in the dynamical properties of the system must be driven by more subtle features of the topology than connectedness. In fact, as we shall see, there is much about the system that cannot be reduced to any single structural characteristic of the underlying graphs—a warning signal for dealing with any more complicated dynamics.

In order to emphasize the role of population structure, the subsequent analysis is restricted to a simplified model of disease spreading in which each element of the population is in one of three states: *susceptible, infected,* or *removed.* At each discrete point in time ($t$), every infected element can infect each one of its neighbors with probability ($p$), the *infectiousness.* Any newly infected elements remain infected for one time step,[33]

---

[33] The time period ($\tau$) for which an infected agent remains infectious can be set to one without loss of generality. The reason is that $p$ and $\tau$ do not vary independently. In fact, the dynamics for any given $\tau$ can be reproduced with $\tau = 1$ merely by rescaling $p$. In other words, there is complete equivalence (in this model) between being exposed to a more infectious disease for a short period of time and a less infectious disease for a long period of time. This feature of the model greatly simplifies its analysis.

after which they are removed permanently from the population (presumably by immunity or death) and so play no further part in proceedings. Hence, if at $t = 0$ a single element is infected (by some external influence), then at some later time, one of two things must have happened:

1. The disease will have run its course and died out, infecting some fraction of the population and leaving the remaining fraction $F_s$ uninfected,

or

2. The whole population will have been infected ($F_s = 0$) in some characteristic time $T$.

The natural question to ask then is whether or not the structure of the population, expressed in terms of the $\alpha$-graph model presented earlier, has any effect on $F_s$ and $T$ of the related system. Furthermore, if so, can the functional forms of $F_s(\phi)$ and $T(\phi)$ be understood in terms of our structural statistics $L(\phi)$ and $C(\phi)$?

*Results.*—The model has two parameters that can vary independently of each other: the infectiousness $p$, which determines the local dynamics; and the fraction of shortcuts $\phi$. We are now in a position to compare some results for an entire range of topologies as a means of answering the two questions stated above within the very narrow context of this specific dynamical system. Figure 11 shows the steady-state fraction of susceptibles $F_s$ versus $p$ for the two extreme values of $\phi$, where graphs of parameters $n = 1,000, k = 10$ have been used to determine the couplings of the system. Three distinct regions are apparent:

1. For $p \lesssim 1/(k - 1) \approx .11$ all topologies (i.e., all values of $\phi$) yield the same result: the disease infects only a negligible fraction of the population before dying out. This is a *trivial* steady state, because nothing happens that can distinguish between different topologies.
2. For $.11 \lesssim p \lesssim .5$ different topologies yield different $F_s$.
3. For $p \gtrsim .5$, all topologies once again yield the same end result, only this time it is a *nontrivial* steady state because the disease has taken over the entire population.

There is nothing more to say about region 1, but regions 2 and 3 deserve some extra attention. Region 2 is confusing: there appears to be some significant relationship between structure and dynamics, but its mechanism is not transparent. Figure 11 suggests that in region 2 the fraction of the population that will become infected with a disease of some specified $p$ depends significantly on $\phi$, and this is shown more explicitly in figure 12 for a particular value of $p$. In epidemiological terms, this dependence on $\phi$ is equivalent to the statement that the impact of a disease depends not just on how infectious it is, but also on the connective topology of the population. This message is not new in the epidemiological literature, es-
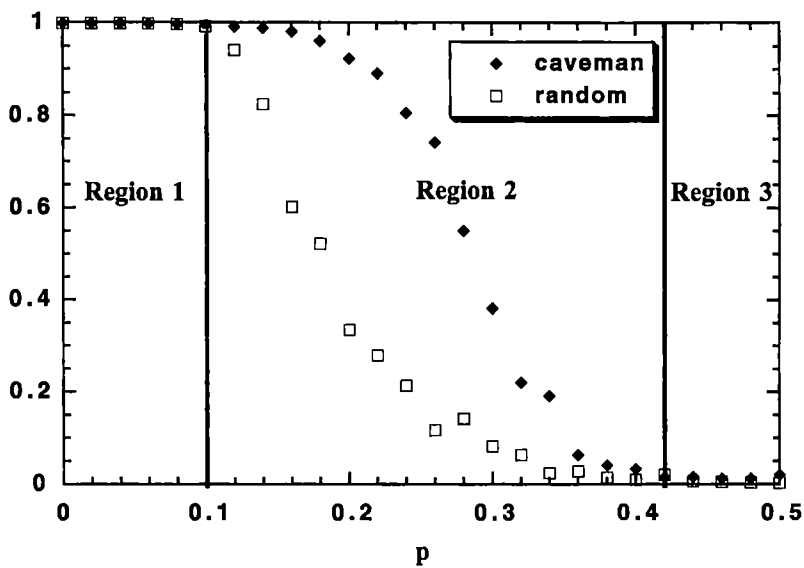
FIG. 11.—Fraction of uninfected survivors ($F_s$) versus infectiousness ($p$) for disease spreading dynamics on a network generated by the $\alpha$-model at clustered and random extremes.
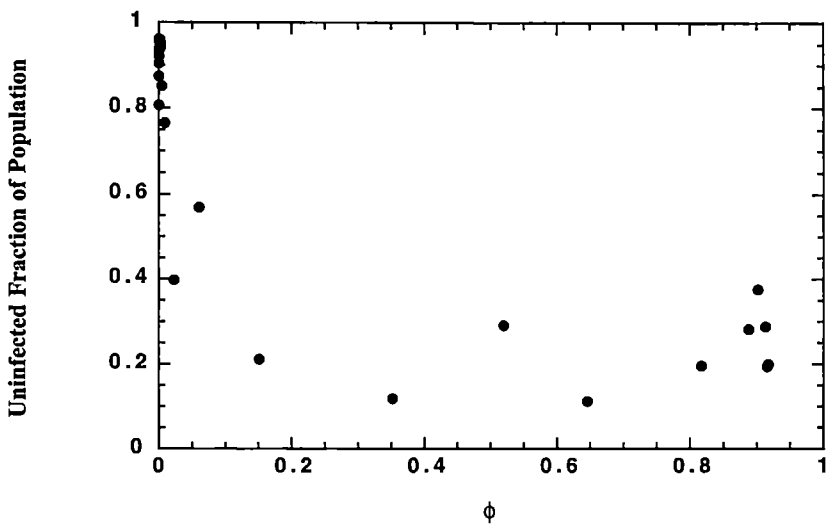


FIG. 12.—$F_s$ versus $\phi$ for $p = .24$
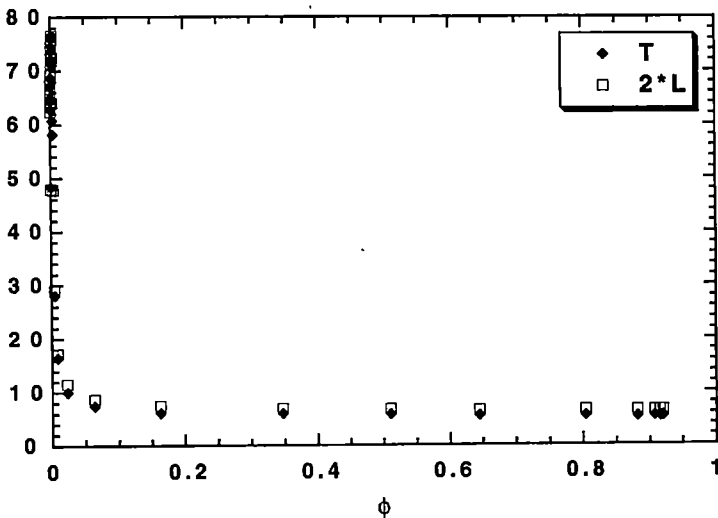
FIG. 13.—Time taken to reach a steady state ($T$) versus fraction of shortcuts ($\phi$) in the underlying $\alpha$-graph for the disease spreading model.

pecially that of sexually transmitted diseases. What is new is the idea that the structural changes required can be very subtle. Unlike the difference between a chain and a star graph or a ring lattice and a random graph, the difference between a big- and a small-world graph can be a matter of only a few randomly rewired edges—a change that is effectively undetectable at the level of individual vertices. Unfortunately, it is difficult to be any more precise than this, as the functional form of $F_s(\phi)$ is not clear. Nevertheless, even this broad-brush observation has implications for the way we think about social diseases, which are often perceived as confined to isolated subgroups of a population. The message here is that the highly clustered nature of small-world graphs can lead to the intuition that a given disease is "far away" when, on the contrary, it is effectively very close.

Region 3 is a simpler matter to examine, as the disease always takes over the entire population regardless of its connective topology. It is merely a question of how long this takes. However, figure 13 shows that in this region the time taken to reach the steady-state $F_s = 0$ varies dramatically as a function of $\phi$ and that disease can spread on a small-world graph far more rapidly than in a connected caveman graph and almost as fast as in a random graph. In fact, figure 13 also shows that when $p = 1$, $T(\phi)$ bears a close functional relationship to $L(\phi)$.[34] Hence, in this parameter

[34] For $p = 1$, this functional similarity might be expected, as $T$ is just the maximum distance from the initially infected vertex to any other vertex in the graph. For a

regime at least, the transient time for the dynamics to reach the same, nontrivial steady state bears a simple and obvious relationship to the structure of the underlying graph. That is, shorter characteristic path length implies faster spreading of the disease. In a real-world scenario, where an epidemic can be responded to, the time scale on which it spreads becomes a crucial factor. A striking consequence of this result is that, in a small-world graph, the characteristic time scale has become very small, but the clustering $C$ is still large. As with the extent of disease spread in region 2, the change in structure in region 3 that causes the disease to spread much faster may not be observable at a local level.

Although the relationship between population structure and disease dynamics is not always clear, even for this simple model, it does appear as if the gross features of the dynamics are dominated by the characteristic path length of the underlying graph. This is by no means universally true of dynamical systems on graphs. For instance, when the transmission of a behavioral trait exacts a cost, such as for a prisoner's dilemma model of cooperation, a high degree of reciprocity may be necessary for the trait to survive in the population (Axelrod 1984; Boyd and Richerson 1988; Cohen, Riolo, and Axelrod 1999). In the case of small groups of cooperators struggling for survival in a sea of defectors, Axelrod (1984) notes that cooperators must interact preferentially, which in network terms is equivalent to high local clustering. In a highly clustered graph, cooperators located in the same cluster can survive—even thrive—in the midst of a noncooperative majority. Conversely, in a random graph (with negligible clustering), any small group of initial cooperators will be eroded from the periphery, as each peripheral cooperator will be interacting predominantly with defectors, thus failing to reap the benefits of reciprocity. As with disease spreading, there is a transition between these two extremes, but in this case, the dynamics tend to be dominated by the clustering coefficient rather than by the characteristic path length.[35]

This result suggests an interesting role for small-world architectures, which by virtue of their short characteristic path length and high cluster-

---

perfect ring structure $T = D$, where $D$ is the diameter of the graph and $D = 2L$ Hence, for even the approximate ring structure generated for an $\alpha$-graph with $\phi = 0$, we could expect that $T \approx 2L$. Furthermore, for *any* graph, it is necessarily true that $L \leq T \leq D$. For random graphs, where the number of vertices at a distance $d$ from any vertex generally grows exponentially (Bollobás 1985), $D < 2L$ so we could also expect that $L \leq T \leq 2L$ for *any* value of $\phi$. Thus $T$ should be related to $L$ through nothing more than a multiplicative factor $1 \leq c(\phi) \leq 2$, which is not significant on the scale of the changes occuring in both statistics as a function of $\phi$. Hence, the two curves in figure 13 might be expected to look very much alike, as they do.

[35] There are a number of subtleties to this result which are explained more completely in Watts (1999, chap. 8).

ing coefficient, can support the rapid dissemination of information without necessarily compromising behavior that is individually costly but beneficial when reciprocated. Furthermore, the more general idea of optimizing an architecture to satisfy two or more opposing constraints may prove a useful concept in the design or modification of large networks such as organizations.

## SUMMARY

This article examines a particular class of graphs that interpolates between highly ordered and highly random limits. A significant feature of these graphs is that the presence of a very small fraction of long-range shortcuts can lead to the coexistence of high local clustering and a small global length scale. The superposition of these apparently contradictory properties is a graph-theoretic formalization of the small-world phenomenon. The motivation for the small-world phenomenon comes from social networks, but it turns out to be a much more general effect that arises under quite weak conditions in large, sparse, partly ordered and partly random networks. Its existence is not predicted by current network theories, yet it seems likely to arise in a wide variety of real networks, especially in social, biological, and technological systems. One consequence of this result is that it is highly likely that the phenomenon exists in the real social world—a notion currently supported by only limited data but consistent with anecdotal experience.

In addition to their interesting structural properties, small-world graphs are also relevant to the social and natural sciences through their effect on the globally emergent features of dynamical systems. Specifically, distributed dynamical systems can exhibit dramatically different behavior on small-world networks—an effect that may have implications in fields as diverse as public health and organizational behavior and design.

## APPENDIX

The algorithm for constructing a graph according to equation (5) proceeds as follows:

1. Fix a vertex $i$.
2. For every other vertex $j$, compute $R_{i,j}$ according to equation (5), with the additional constraint that $R_{i,j} = 0$ if $i$ and $j$ are already connected.
3. Sum the $R_{i,j}$ over all $j$, and normalize each to obtain variables $P_{i,j} = R_{i,j}/(\sum_{l \neq i} R_{i,l})$. Then, since $\sum_j P_{i,j} = 1$, we can interpret $P_{i,j}$ as the probability that $i$ will connect to $j$. Furthermore, we can interpret $P_{i,j}$ geometrically as follows: divide the unit interval $(0, 1)$ into $n - 1$ half-open subintervals with length $P_{i,j} \ \forall \ j \neq i$.

4. A uniform random variable is then generated on (0, 1). It must fall into one of the subintervals, say the one corresponding to $j_*$.
5. Connect $i$ to $j_*$.

This procedure is then repeated until the predetermined number of edges ($M = kn/2$) has been constructed. The vertices $i$ are chosen in random order, but once a vertex has been allowed to "choose" a new neighbor, it may not choose again until all other vertices have taken their turn. However, vertices may be "chosen" arbitrarily often, and this leads to a nonzero variance in the degree $k$. But the fact that all vertices are forced to make one new connection before any others are allowed to choose a second time ensures at least that no vertices will be isolated (as long as $k \geq 2$).

REFERENCES

Achacoso, T. B., and W. S. Yamamoto. 1992. *AY's Neuroanatomy of C. elegans for Computation*. Boca Raton, Fla.: CRC Press.

Arthur, W. Brian, and David Lane. 1993. "Information Contagion." *Structural Change and Economic Dynamics* 4 (1): 81–104.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Bollobás, Bela. 1985. *Random Graphs*. London: Academic Press.

Bollobás, Bela, and Fan R. K. Chung. 1988. "The Diameter of a Cycle Plus a Random Matching." *SIAM Journal of Discrete Mathematics* 1 (3): 328–33.

Boyd, Robert, and Peter J. Richerson. 1988. "The Evolution of Reciprocity in Sizable Groups." *Journal of Theoretical Biology* 132:337–56.

Burt, Ronald S. 1982. *Toward a Structural Theory of Action*. New York: Academic Press.

———. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, Mass.: Harvard University Press.

Cerf, V. G., D. D. Cowan, R. C. Mullin, and R. G. Stanton. 1974. "A Lower Bound on the Average Shortest Path Length in Regular Graphs." *Networks* 4:335–42.

Cohen, Michael D., Rick L. Riolo, and Robert Axelrod. 1999. "The Emergence of Social Organization in the Prisoner's Dilemma: How Context-Preservation and Other Factors Promote Cooperation." Working Paper 99-01-002. Santa Fe Institute.

Coleman, James S. 1973. *The Mathematics of Collective Action*. Chicago: Aldine.

Chung, Fan R. K. 1986. "Diameters of Communications Networks." Pp. 1–18 in *Mathematics of Information Processing*, vol. 34 in Proceedings of Symposia in Applied Mathematics. Providence, R.I.: American Mathematical Society.

Crick, Francis, and Christof Koch. 1998. "Constraints on Cortical and Thalamic Projections: The No-Strong-Loops Hypothesis." *Nature* 391:245–50.

Davidson, Mark L. 1983. *Multidimensional Scaling*. New York: Wiley.

Davis, James A. 1967. "Clustering and Structural Balance in Graphs." *Human Relations* 20:181–87.

Fararo, Thomas J., and M. Sunshine. 1964. *A Study of a Biased Friendship Net*. Syracuse, N.Y.: Syracuse University Press.

Foster, Caxton C., Anatol Rapoport, and Carol J. Orwant. 1963. "A Study of Large Sociogram: Estimation of Free Parameters." *Behavioral Science* 8:56–65.

Freeman, Linton C. 1979. "Centrality in Social Networks: Conceptual Clarification." *Social Networks* 1:215–39.

———. 1982. "Centered Graphs and the Structure of Ego Networks." *Mathematical Social Sciences* 3:291–304.

Friedkin, Noah E. 1990. "Social Networks in Structural Equation Models." *Social Psychology Quarterly* 53 (4): 316–28.

———. 1991. "Theoretical Foundations for Centrality Measures." *American Journal of Sociology* 96 (6): 1478–1504.

Gladwell, Malcolm. 1996. "The Tipping Point." *New Yorker,* June 3, pp. 32–38.

Granovetter, Mark S. 1973. "The Strength of Weak Ties " *American Journal of Sociology* 78 (6): 1360–80.

———. 1974. *Getting a Job: A Study of Contacts and Careers.* Cambridge, Mass : Harvard University Press.

———. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83 (6): 1420–43.

Grossman, Jerry W., and Patrick D. F. Ion. 1995. "On a Portion of the Well-Known Collaboration Graph." *Congressus Numeratium* 108:129–31.

Guare, John. 1990. *Six Degrees of Separation: A Play.* New York. Vintage Books.

Harary, Frank. 1959. "Status and Contrastatus." *Sociometry* 22:23–43.

Hassell, Michael P., Hugh N. Comins, and Robert M. May. 1994. "Species Coexistence and Self-Organizing Spatial Dynamics." *Nature* 370:290–92.

Herz, Andreas V. 1994. "Collective Phenomena in Spatially-Extended Evolutionary Games." *Journal of Theoretical Biology* 169:65–87.

Hess, George. 1996a. "Disease in Metapopulation Models: Implications for Conservation." *Ecology* 77 (5): 1617–32.

———. 1996b. "Linking Extinction to Connectivity and Habitat Destruction in Metapopulation Models." *American Naturalist* 148 (1): 226–36.

Kareiva, Peter. 1990. "Population Dynamics in Spatially Complex Environments: Theory and Data." *Philosophical Transactions of the Royal Society of London* 330: 175–90.

Kauffman, Stuart A. 1969. "Metabolic Stability and Epigenesis in Randomly Constructed Genetic Nets." *Journal of Theoretical Biology* 22:437–67.

Kochen, Manfred. 1989. "Toward Structural Sociodynamics." Pp. 52–64 in *The Small World,* edited by Manfred Kochen. Norwood N.J.: Ablex.

Korte, Charles, and Stanley Milgram. 1970. "Acquaintance Linking between White and Negro Populations: Application of the Small World Problem." *Journal of Personality and Social Psychology* 15:101–18.

Kretzschmar, Mirjam, and Martina Morris. 1996. "Measures of Concurrency in Networks and the Spread of Infectious Disease." *Mathematical Biosciences* 133:165–95.

Longini, Ira M., Jr. 1988. "A Mathematical Model for Predicting the Geographic Spread of New Infectious Agents." *Mathematical Biosciences* 90:367–83.

March, L , and P. Steadman. 1971. Chapter 14 in *The Geometry of Environment.* London: RIBA Publications.

May, Robert M. 1995 "Necessity and Chance: Deterministic Chaos in Ecology and Evolution." *Bulletin of the American Mathematical Society* 32:291–308.

Milgram, Stanley. 1967. "The Small World Problem." *Psychology Today* 2:60–67.

Murray, James D. 1991. *Mathematical Biology,* 2d ed. Berlin: Springer Verlag.

———. 1993. *Mathematical Biology,* 2d ed. Berlin: Springer Verlag.

Nowak, Martin A., and Robert M. May. 1993. "The Spatial Dilemmas of Evolution." *International Journal of Bifurcations and Chaos* 3 (1): 35–78.

Phadke, Arun G., and James S. Thorp. 1988. *Computer Relaying for Power Systems.* New York: John Wiley.

Pool, Ithiel de Sola, and Manfred Kochen. 1978. "Contacts and Influence." *Social Networks* 1:1–48.

Rapoport, Anatol. 1953a. "Spread of Information through a Population with Socio-

Structural Bias: I. Assumption of Transitivity." *Bulletin of Mathematical Biophysics* 15:523–33.

———. 1953*b*. "Spread of Information through a Population with Socio-Structural Bias: II. Various Model with Partial Transitivity." *Bulletin of Mathematical Biophysics* 15:535–46.

———. 1957. "A Contribution to the Theory of Random and Biased Nets." *Bulletin of Mathematical Biophysics* 19:257–71.

———. 1963. "Mathematical Models of Social Interaction." Pp. 493–579 in *Handbook of Mathematical Psychology,* vol 2 Edited by R. R. Bush, R. D. Luce, and E. Galanter. New York: Wiley.

Rouvray, Dennis H. 1986. "Predicting Chemistry from Topology." *Scientific American* (September): 40–47.

Sattenspiel, Lisa, and Carl P. Simon. 1988. "The Spread and Persistence of Infectious Diseases in Structured Populations." *Mathematical Biosciences* 90:341–66.

Skvoretz, John. 1985. "Random and Biased Networks: Simulations and Approximations." *Social Networks* 7:225–61.

Solmonoff, Ray, and Anatol Rapoport. 1951. "Connectivity of Random Nets." *Bulletin of Mathematical Biophysics* 13:107–17.

Strogatz, Steven H , and Ian Stewart. 1993. "Coupled Oscillators and Biological Synchronization." *Scientific American* 269 (6): 102–9.

Tjaden, Brett. 1997. E-mail to author, May 14.

Watts, D. J. 1999. *Small Worlds: The Dynamics of Networks between Order and Randomness.* Princeton, N.J.: Princeton University Press.

White, Harrison. 1970. "Search Parameters for the Small World Problem." *Social Forces* 49:259–64.

White, J., E. Southgate, J. N. Thompson, and S. Brenner. 1986. "The Structure of the Nervous System of the Nematode *Caenorhabditis elegans."Philosophical Transactions of the Royal Society of London,* Series B, 314:1–340.

Wiener, Harry. 1947. "Structural Determination of Paraffin Boiling Points." *Journal of the American Chemistry Society* 69:17–20.

Yamaguchi, Kazuo. 1994*a*. "The Flow of Information through Social Networks Diagonal-Free Measures of Inefficiency and the Structural Determinants of Inefficiency." *Social Networks* 16:57–86.

———. 1994*b*. "Some Accelerated Failure-Time Regressive Models Derived from Diffusion Process Models: An Application to Network Diffusion Analysis." *Sociological Methodology* 24:267–300.

# Book Reviews

*The Sociology of Philosophies: A Global Theory of Intellectual Change.*
By Randall Collins. Cambridge, Mass.: Harvard University Press, 1998.
Pp. xix+1098. $49.95.

Andrew Abbott
*University of Chicago*

This book is the large work many of us have expected for years from
Randall Collins. It sets forth a theory of intellectual development which
it then imposes on case histories from widely scattered traditions. After
a brief introduction, there are two chapters of pure theory, one chapter of
detailed illustration of the theory, four analytic histories of Asian thought
systems, seven somewhat more detailed analytic discussions of moments
in European philosophies, plus two chapters of reflection. The work
shows breathtaking ambition.

Collins says he has two aims: first to set an example by writing world
history, and second to offer a "causal explanation" of ideas via the concept
of interaction networks. There is little question that he has succeeded in
the first aim. The book is awesome in scope and literacy. One can only
applaud the venture; Collins is surely right to think that in the coming
century Western academics should know the names and ideas of Chuang
Tzu and Saicho as they do those of Aquinas and Machiavelli.

With respect to the aim of explaining intellectual life, the record is more
mixed. One's most optimistic feeling after reading the purely theoretical
chapters is that Collins's heart is not in them. The theory is a combination
of Durkheimianism—both micro and macro—with a transactional analy-
sis that borrows most of its language and conceptions from simple ratio-
nal action theory. Collins accepts (to my eye, uncritically) both the social
massivism of Durkheim and the various radical simplifications of the lan-
guage of payoffs, optimization, and deliberation. Alternative frameworks
are not seriously considered. There are gestures in the direction of serious
structuralism (*à la* Simmel or Park or Harrison White) but Collins is un-
willing to relax his Durkheimian core; he just wants to loosen it a little.

Most important, for a book that will be about the mutabilities of mean-
ing, *The Sociology of Philosophies* accepts an extraordinarily wooden no-
tion of culture. Absent from a bibliography that clearly intends com-
prehensiveness are Bakhtin, Evans-Pritchard, James, Jakobson, Peirce,
Sapir, Saussure, and Whorf. The lability of language, the multiplicity of
meanings, the inherent indexicality of all interaction and experience:
these are completely absent here, an absence that parallels, at another
level, the relative disattention (in the theoretical argument) to the actual

content of the philosophies to be theorized. Their social structure gets all the attention, and in considering that structure Collins assumes throughout the theoretical chapters that hierarchies of meaning and stratifications of ideas are both clear at a time and more or less stable over time. These chapters are also marred by stylistic and intellectual unevenness. Strong but uncontroversial assertions are intermingled with interesting asides, radical statements, simple errors, and penetrating insights.

In the short theoretical/empirical chapter on networks we see that Collins takes his methodology very seriously indeed. Here he measures the importance of philosophers by the proportion of pages mentioning them in various downstream sources. He believes that relative percentages of pages devoted in a later present to a particular ancient philosopher measure something direct about the long-term productivity of that person's work. Although I cannot find any real discussion of the matter, "connection" between philosophers (the foundation of the many network diagrams in the book) seems to be measured by the interpretations of standard secondary sources. The dozens of methodological, historical, and theoretical assumptions involved in all this get no real consideration. Nor is there recognition of the extensive formal modeling of such systems by Cavalli-Sforza and Feldman.

To my mind, such an analysis is justified only by a combination of blunt honesty about the intellectual costs involved with a bravura and energy that creates in readers the necessary suspension of disbelief. Neither the honesty nor the bravura are here. The reader sees objections long before Collins comes to them. There is in fact no systematic discussion of measurement. Collins never presents a formal model for the chain that connects a philosopher's original writing through all the various retellings to survival in later sources. Nor does he really consider alternative theories. He never considers that the restriction of important philosophers to a particular few might inhere in human hardware restrictions, or that it perhaps might arise in the rhetorical structure of historical narrative. Or that present fame might arise because a person's work is "good to think with" in the present context, or that present reputation might simply represent convention. One has only to teach a class in some classical tradition (or to listen to a few papers by graduate students and colleagues) to realize that many very generally available ideas are attributed to particular people, a fact that vitiates the entire methodology. Collins's basic answer to this challenge, and to most others (see, e.g., appendix B), is that everything comes out in the wash and that in any case the uncertainty of downstream measurement "is not a limitation on what we can understand of intellectual creativity; it is the nature of the beast." That is, later pages *are* influence. I do not find this position credible.

The same argument applies to Collins's contention that well-known philosophers have the largest number of ties to other known philosophers. This result (whether we accept it is another matter) could also be generated by various artifactual processes: differential survival about information on connections (we are more likely to know about lesser people be-

cause of their connections with a particular major figure), rhetorical necessity (a story is more lively if its protagonists are seen to be in personal conflict—this makes such information more likely to survive), and so on. Collins talks about certain writers as overrated (e.g., Confucious) but still thinks his measures work on balance. This reader does not believe him.

The theoretical and methodological sections, even in the most optimistic view, do not justify trusting Collins's interpretations of the thousands of empirical sources that he has seen and we the readers have not. We do not believe him capable of finding in those sources a rejection of his argument. Like many other readers, I sought here and there diagrams covering people I do know and always found minor things I could dispute (e.g., on p. 328 a positive acquaintance link is shown between the great Japanese writers Sei Shonagon and Murasaki Shikibu, who in fact not very cordially hated each other). I compare the diagram on page 96 with that on page 117 and wonder whether, if Collins had not provided the labels, I would have seen "recombination of schools" in one and "syncretisms" in the other. I have, indeed, no belief in the replicability of any of the details of the analysis and hence no belief that they really support the main theoretical contentions nor, given Collins's ex ante belief in his hypothesis, even a belief that they are viable mini histories of the philosophical epochs involved.

I do think the general point of the book—that intellectual traditions tend to happen in limited numbers of schools—is probably right. A serious analysis of the possible theoretical reasons for this limitation would have been of great interest. But it is not here. For Collins is not really passionate about the theory, not insistent that he (and we) get it right, not insistent that he foresee and reject all possible objections. His heart is really in the project of empirical comprehensiveness, the getting of everything known about knowledge onto a single grand template. He has the key to all mythologies.

Even the writing is, at best, not beguiling. Collins's style comprises long sentences, ponderous tone, and endless repetition. The very length militates against lightness, although Braudel's *Mediterranean* shows one can write a monumental book without somnolence. Finally, a note about copyediting. The Harvard University Press ought to be able to make sure subjects and verbs agree in number (e.g., on p. 70: "In my opinion they do not shake the theoretical results, and even reinforces them.")

Buried in this book are some great insights. The labor it embodies is immense. The underlying argument, were it reduced to a short and powerful theoretical analysis, would be an important contribution. But as it is, I cannot recommend this book.

*Analytic Narratives.* By Robert H. Bates, Avner Greif, Margaret Levi, Jean-Laurent Rosenthal, and Barry R. Weingast. Princeton, N.J.: Princeton University Press, 1998. Pp. x+249. $50.00 (cloth); $19.95 (paper).

Jack A. Goldstone
*University of California, Davis*

It is remarkable that a book by such brilliant scholars can be so uneven. The book is nothing if not ambitious—its authors repeatedly claim that their method, which they label "analytic narrative," allows deep and general insights that have escaped historians. The method is straightforward—construct a narrative of critical events or outcomes from the historical literature and then show how those events or outcomes could have developed as equilibria solutions to formal games. Why bother with the formal games? If one can construct a model in which the actual historical event is the unique outcome of simple initial conditions and rules of behavior, then we gain insights into why that particular outcome, and not others, resulted in real-world history.

But this ideal proves difficult in practice. First, the construction of the accounts is a highly debatable enterprise. Nor are the models always helpful. As game theorists know all too well, most games have many, many equilibria solutions. The one that actually develops depends on initial conditions and the path of play. So even with these game models, the authors are often reduced to saying that particular conditions produced a particular outcome (no surprise to historians) while still arguing that general principles of self-interest and strategic behavior motivate this particularistic account.

Avner Grief's chapter on the political structure of late medieval Genoa is reminiscent of John Padgett's analysis of Florence. In both cases, particular outcomes are shown to be the result of strategic maneuvers by noble clans. Grief develops the thesis that Genoa needed to solve a specific maximization problem: how to let the economy grow while limiting destructive efforts by one clan to claim the whole pot. Initially, the problem was solved by limiting Genoa's overseas possessions so that the pot was not large enough to tempt any clan into a struggle to seize assets. But as the pot grew larger, the prize grew too tempting, and civil war resulted. Or rather it should have, had the threat posed by Frederick Barbarossa from 1154 to 1164 not led the clans to unite to defend Genoa. Thus history's incidentals trump the model for a decade, and when civil war does follow Barbarossa's withdrawal, peace is finally attained with the appointment of a foreign governor. An equilibrium solution to halt conflict? Yes, but not one that was found by the Genoese; the foreign governor was the idea of Henry IV, who installed such representatives throughout Italy. The equilibrium model does, however, help explain why the Genoese adopted the institution of foreign governors and appointed them for over another century.

Jean-Laurent Rosenthal then offers insights into the tax and finance systems of prerevolutionary France. While Rosenthal knows the financial markets well, he starts from what seems to me a premise only an economist could love—France should have aimed at "efficient" taxation, vesting all tax authority either in the king or in a parliament, because that would have been better for the nation. Now in the real world, it was the goal of all enlightened regime builders of the 18th century, in the United States, England, and France to *avoid at all costs* an "efficient" tax regime, because such a concentration of unopposed power would threaten liberties. Indeed, the Federalist papers and the U.S. Constitution are filled with quite deliberate inefficiencies explicitly intended to preserve liberty. So when Rosenthal argues that England had an efficient tax system, while France did not, and bends over backward to find institutional reasons for this difference, the whole argument seems to me to go awry. Rosenthal is right in one respect—England did collect far more taxes per capita that France. But consider that, circa 1700, France had roughly four times the population of England, yet the two powers were engaged in naval and land wars for decades. So England faced a very straightforward dilemma—either raise several times as much money per capita as the French or be overwhelmed. Such a situation wonderfully concentrates the mind, and it was not exclusive royal or parliamentary power that led to such high taxation but provisional agreements between crown and elites to save the kingdom. Napoleon did produce "efficient" taxation, and a decade of civil war and eventual catastrophic military defeat. Sixty years later, facing a revolution in America assisted by France, it was the English who failed. Divisions between crown and parliament continued to trouble England until the late 19th century.

Barry Weingast offers an interesting study of how sectional interests in the 19th-century United States led to the Civil War. He highlights the "balance rule": for every free state admitted to the Union, a slave state should also be admitted to preserve balance in the Senate as the key to stability. Weingast chides historians for arguing over the importance of local politics versus slavery in the run-up to the Civil War, pointing out that the "balance rule" was key because it provided the only "stable equilibrium" for national politics between opposing regional interests. As the Union expanded north and west into areas where plantation slavery was impractical, the balance rule could not be maintained, and the South had to secede or lose on all regional issues. Fair enough, but hardly that new or interesting; putting some of the more extreme polemics aside, historians worry about local politics to fill in the picture on many levels, not because they doubt the importance of the balance rule and the debate over slavery. I am reminded of studies of the English Revolution, where for some years scholars tried to depict the event as the product of local conflicts and not national political issues; serious scholarship today sees the English Revolution as involving the interplay of local cleavages with national concerns about royal power and taxation. Similarly, the Civil War interwove local concerns with national policies, and nothing is

gained by focusing so exclusively on a single "principle" to explain stability or conflict.

The strongest chapters of this book are those that do not seek a single game solution as the explanation for a complex historical problem. Margaret Levi's analysis of how conscription developed over time and across countries, with important aspects such as commutation and substitution being ended at different times in different countries, uses models of interests to illuminate different trajectories. Her conclusion that "norms of fairness and equity" need to be integrated with interests to understand outcomes is a major step forward in the application of analytic models to history. Robert Bates's wonderful chapter surprises in that he tries and *rejects* a number of formal analytic models. Unlike the other chapters, for Bates, when history contradicts the model, so much the worse for the model. Bates's study of the international coffee trade leads to a remarkable conclusion: the cartel of coffee producers was able to last only because it was supported by a major coffee consumer, namely the United States. During the Cold War, the United States wanted to support the economies of developing countries, especially in Latin America, and so accepted their export regimes. Bates cleverly shows how even the major coffee roasters in the United States adapted to the regime of cartelization of coffee bean production, using their relations with the cartel to gain control of the U.S. consumer market. In the end, the cartel that appears adversarial to coffee roasters and consumers was in fact supported precisely by them.

On the whole, this book does a good job of illustrating the gains (and the pitfalls) of using formal game models to examine historical events. At worst, it shows how economists' tendencies to frame everything in terms of seeking optimal and equilibrium outcomes can distort a history in which many actors have no sense of what is optimal and in which random interactions constantly undermine equilibria. But at the same time, judicious use of such models can make it clear where actors depart from simply expected outcomes and can guide the search for shifts in norms, or for additional influences (e.g., the Cold War strategic concerns) that intrude on simple producer-consumer interactions. As a method, analytic narrative is clearly still in its infancy, but it has promise and bears watching.

*Durkheim and the Jews of France.* By Ivan Strenski. Chicago: University of Chicago Press, 1997. Pp. ix+215. $41.00 (cloth); $15.90 (paper).

Pierre Birnbaum
*Université de Paris*

As we celebrate the centennials of Durkheim's successive works, and as the knowledge of Durkheimian sociology as well as of French Jews of the end of last century is making great progress, one can legitimately

examine the relationship between Durkheim and French Judaism. There already exists a well-known and abundant literature on this topic, but a systematic treatment is certainly justified. It is also wise, as Strenski suggests, to refrain from any essentialist vision presupposing an eternal Jewish soul: "Essentialist treatments of Durkheim's relation to things Jewish or to Judaism simply fail to meet criteria of what would count as historical assertions about Durkheim" (p. 3). This firm rejection of an essentialist perspective, which one can only approve of, is even more justified in the author's opinion, as "Durkheim disavowed Judaism as irrelevant. . . . Throughout nearly all of his life, Durkheim seemed to have resisted identification of himself as Jewish. . . . He was embarrassed—even among Jews in his hometown—to be singled out as Jewish. . . . Durkheim was not in effect a modern Marrano" (pp. 4–5). Without demonstrating such assertions, Strenski believes that Durkheim is simply being sensitive, and only at the end of his life, to anti-Semitism. The book's novel thesis nevertheless consists in establishing an intimate link between the works of scholars belonging to the *Science du Judaisme,* who rejected interpretations hostile to the period's dominant Jewish traditions (Frazer, Renan, etc.), and the research of Henri Hubert and Marcel Mauss on the essential functions of ritual and religious institutions, hence the renaissance of an implicit Jewish identity leading to a theory of religion as social ritual. In this sense, the theory of religion sustained by the entire Durkheimian team would be based on a defense of Judaism attacked by anti-Semitic Catholic scholars of the time. Hubert and Mauss, Durkheim's closest collaborators, would have been formed by the last representatives of the *École de la Science du Judaisme,* the French version of the *Wissenschaft des Judentums:* Hartwig Derenbourg (Islamist), Israël Levi (Talmudist), or Sylvain Lévi (Indologist) taught Hubert and Mauss "and thus eventually Durkheim himself" (p. 90). This is where Durkheim's latent and necessarily vague Judaism could be found. Because he cannot marshal evidence to account for it, Strenski constantly uses purely hypothetical expressions in order to establish eventual connections between scholars. For example, Strenski supposes an intellectual connection between Salomon Munk coming to France around 1840, Levi, Mauss, and Durkheim. Simply because Munk influenced Sylvain Lévi, who taught Mauss at the École des Hautes Etudes, and Mauss later because Durkheim's closest collaborator. For Strenski, "Munk may very well have indirectly shaped the Durkheimian approach to the study of religion" (p. 95). In Strenski's genealogy, these links are forged too easily, and one might hesitate before trusting the resulting chain to bear weight.

If it is legitimate to trace back intellectual influences, the trajectories of ones and of others, the risk of falling into a straightforward history of ideas without great sociological rigor is a real one. By rightly rejecting an essentialist perspective, Strenski falls here into unprovable, evanescent collaborations, merely establishing a connection between Hubert, Mauss, and scholars of Judaism and without getting with certitude to Durkheim

himself other than by supposition. Oddly, this work concerns, in the end, much more Mauss than Durkheim. Highlighting certain influences on Mauss, the author nevertheless remains almost mute on many aspects of Mauss's intellectual formation, or, more specifically, on his relationship with a socialism that deliberately distances him, according to Mauss himself, from Judaism. Strangely, this book remains silent on Durkheim and only evokes a few names from the period's intellectual Judaism, neglecting altogether "the Jews of France." The book's title is thus awkward. It is also odd that the author does not take into account the most recent works on these questions. Strenski cites texts outside the topic itself, dating from 1996, but overlooks numerous works that also reject any essentialist dimension but that nevertheless are different from classical conceptions, such as that of Steven Lukes (whom Strenski favorably cites on this point). He thereby reduces any influence of Judaism on Durkheim's thought. He would have in this way avoided confining himself to a highly debatable conception of assimilation as entirely negating identity. One is thus astonished to see that books by Perrine Nahum-Simon, Aron Rodrigue, and Michael Graetz, as well as the fundamental work of Marcel Fournier and many other works on Durkheim and on French Jews of the period (by Béatrice Philippe, Philippe Landau, etc.) were not used. One cannot understand that, on such a topic, the Dreyfus Affair is merely evoked when it was of great concern for Jews of France and especially for Durkheim, who engaged himself with passion in the *Ligue des droits de l'homme* and personally confronted, with courage, anti-Semitic attacks in Bordeaux, at the very moment when he was embarking himself in the writing of his essential works. One is surprised to see that Durkheim's numerous letters regarding the Dreyfus Affair and anti-Semitism are never used or cited, just as recent works on the subject related to the affair's commemoration in 1994 are ignored. Last, it is hard to conceive that outside a brief allusion, for example, to *Suicide* (p. 17), Strenski systematically overlooks the numerous passages of the works in which Durkheim devoted himself to long analyses of Jews. One wonders about the quasi-absolute silence of Strenski regarding Durkheim's numerous personal signs of attachment to Judaism throughout his life, from the very solemn respect of certain Jewish holidays to his burial in the cemetery's Jewish square, his tombstone bearing a Hebraic inscription.

In this sense, in spite of interesting connections that fall within an adventurous history of ideas, this book, in which so many names of French or American authors are misspelled, remains empty of any new empirical analysis, even though it was possible to consult many still-unpublished archival sources (e.g., at the Collège de France, l'École Pratique des Hautes Études, and Salomon Reinach's archives). Ultimately, the book does not really treat Durkheim or the French Jews, and even less their relationship.

*Translated by Geneviève Zubrzycki*

*Anthony Giddens: The Last Modernist.* By Stjepan G. Mestrovic. New York: Routledge, 1998. Pp. x+242. $75.00 (cloth); $24.99 (paper).

Charles Camic
*University of Wisconsin—Madison*

Over the course of the past 20 years, the writings of Anthony Giddens have emerged as one of the most substantial bodies of work in the area of contemporary sociological theory. The subject of frequent discussion and debate, Giddens's work has already called forth a sizable literature of exegesis, analysis, and critique. The need for serious further contributions to this literature, however, is widely recognized.

Stjepan Mestrovic's *Anthony Giddens: The Last Modernist* falls short of meeting this need. Thin in exegesis, tendentious in analysis, and querulous in critique, his scattershot and repetitive commentary adds little to the growing scholarship on Giddens's work. The reason for this is not hard to seek. Mestrovic is plainly frustrated—frustrated at contemporary sociology as a whole for everything from its overreliance on "the deductive methodologies of the natural sciences" (p. 70), to its valorization of "mainstream journals" (p. 198), to its practice on holding its annual conventions "in luxury hotels in large metropolitan centers" (p. 216). Mestrovic then projects this gnawing frustration onto Giddens, casting him as the "representative of what [is] wrong with modern sociology" (p. 1). Thrown into this unenviable part, Giddens is doomed from the outset of the book. In Mestrovic's eyes, his writings are not only "shallow," "arrogant," "trite and superficial," "little more than rhetoric, cliches, and slogans" (pp. 11, 173, 194, 205), they are also "symptoms of dissolution and apocalypse," containing the "potential for a new form of totalitarianism" (pp. 161, 212). Indeed, Mestrovic goes so far as to state, "I maintain that, as of this writing, the West's complicity in the genocide in Bosnia is the most serious indictment of the goals and values that Giddens holds dear" (p. 213).

But whatever one's opinions about Giddens's contribution, surely he deserves better than this. Commentators, even critical commentators, have an important obligation to their subjects. As Mestrovic himself recognizes (in the context of still another criticism of Giddens), they must try "to understand [their subjects] with some degree of empathy on [their] own terms" (p. 148). When it comes to the interpretation of Giddens, however, Mestrovic violates this rule, offering "not a sympathetic reading, [but one that] is polemical" (p. 1)—so polemical that Mestrovic fails even to hear some of Giddens's plainer statements, let alone to understand empathically his overall theory on its own terms (the obvious precondition of any serious criticism). Thus, where Giddens has insisted that social structure is both enabling and constraining, Mestrovic charges him with seeing only enablement (p. 193); where Giddens has drawn freely from other disciplines and welcomed a loosening of sociology's boundaries, Mestrovic faults him for advocating "neat and tidy divisions between so-

ciology and the other social sciences" (p. 209); and where Giddens has held that "agency" is an aspect of all human conduct, Mestrovic hears him as "restrict[ing] agency to modern societies" (p. 206). The list could be extended.

Nor is Giddens the only casualty of Mestrovic's inaccurate readings. Placing Giddens in a lineage running from Comte to Parsons, Mestrovic misreads these (supposed) ancestors as well. In his telling, Comte is an Enlightenment "rationalist" who favored a "deductive," "value-free science," and Parsons is a thinker who slighted nonrational action and "made no place for the category of the sacred" (pp. 50, 71, 77, 97). Each of these claims is almost the mirror opposite of the position actually taken by Comte and Parsons.

Mestrovic organizes his book around some of the standard topics that one expects to find in a study of Giddens, namely, Giddens's method, theory of agency and structure, political sociology, and views on modernity. In covering these, Mestrovic is perhaps most interesting and original on the subject of human agency. For the most part, though, his analysis is less concerned with unpacking specific issues than with establishing two overarching arguments.

The first of these is the thesis that, like many other sociologists, Giddens truncates sociology's intellectual heritage, neglecting thinkers such as Schopenhauer, Toynbee, Spengler, Lazarus, Steinthal, and Wundt, and then "amputating the legacies of Durkheim, Weber, Marx, Simmel, [and] Veblen," among many others (p. 208). But this objection is a flaccid one. Every theoretical work is necessarily selective, including some previous thinkers and ideas and omitting others; Mestrovic operates this way himself, overlooking numerous past authors, while describing those he does embrace—Durkheim in particular—in ways so highly selective as to qualify as "amputation" by his own standard. But tactics of this sort are only seriously problematic when they vitiate the particular theoretical statements that they are used to advance or when they offer themselves as valid *historical* claims about sociology's past. Mestrovic, however, shows neither problem to obtain in Giddens's writings. It will be observed, though, that the latter difficulty actually does characterize Mestrovic's own study, which presents itself as a kind of history of sociology but exhibits scant knowledge of the subject, offering such large misstatements as the notion that "from its inception until the rise of Anthony Giddens's reputation . . . sociology was a wild discipline that was kept alive mostly by amateurs or non-sociologists" (p. 209).

Mestrovic's second major argument targets what he calls Giddens's "hidden modernist agenda," or "modernism lite" (pp. 39, 148). The terms refer to Giddens's attempt to "have his cake and eat it too, [by recognizing that] modernity is problematic, but [subscribing] to it nonetheless"—neglecting, in the process, "that contemporary Western societies exhibit traditional, modern, as well as postmodern characteristics simultaneously in addition to hybrids of these ingredients" (pp. 158, 165). Giddens's ideas about modern society certainly merit scrutiny, and Mestrovic is right to

bring them into the spotlight. His move in this direction yields less than it might, however, due once again to his tendency to mischaracterize Giddens, or even to distinguish his conception of modernity, which embraces the whole postfeudal period, from more restrictive everyday conceptions. Nor do Mestrovic's sweeping assertions about "hybrid" societies take into account the large social scientific literature that long ago reworked the tradition-modernity dichotomy, drawing (as Mestrovic does not) on extensive comparative-historical research.

These lapses accord with an unwelcome note of hubris that echoes through the book, as Mestrovic compares himself to Arthur Schopenhauer battling "the great charlatan," G. W. F. Hegel: "For me to go against Giddens is something like Schopenhauer daring to criticize Hegel. . . . It took a half century for people to finally hear Schopenhauer as a spokesperson for their most hidden thoughts and to finally hear Hegel as someone who had deceived them. . . . Whether or when I will be heard remains an open question" (p. 11). Enough said.

*The Science of Sacrifice: American Literature and Modern Social Theory.* By Susan L. Mizruchi. Princeton, N.J.: Princeton University Press, 1998. Pp. ix+436. $65.00 (cloth); $19.95 (paper).

Alan Sica
*Pennsylvania State University*

In a more ideal academic world than one is likely to know, this fine book could serve as the keystone to a bracing honors course, the syllabus for which would include Herman Melville's *Billy Budd, Sailor,* Henry James's *The Awkward Age,* W. E. B. DuBois's *The Souls of Black Folk,* plus Simmel's essays on individuality, Weber's *Protestant Ethic,* and E. A. Ross's *Sin and Society.* The literary scholar Susan Mizruchi is able to connect these apparently disparate texts by recalling theories of sacrificial acts that intrigued Durkheim and most other sociologists of the classic era, lending their societal analyses a special depth often lacking today. (René Girard is the rare exception, as noted early in the book.) The principal drawback, though, to offering such a course is the resulting victimization its participants would face, for because of it they would become too interesting and thoughtful to fulfil their likely roles in today's cultural environment. In a way that Mizruchi might find ironically satisfying, such students would have become sacrificial lambs on behalf of High Culture upon the altar of workplace cretinism, organized and presided over by mass cultural norms and sensibilities. Nevertheless, if they survived into maturity, such a course might be one of only several worth remembering from their "carefree" days.

Mizruchi is a literary critic and intellectual historian whose book demonstrates that the most innovative and thickly constructed interdisciplinary writing probably relies upon intergenerational transmission of cul-

tural capital in order to attain liftoff speed. Crediting her father the sociologist Ephraim, with loaning her pertinent books in the history of social and cultural theory, she applies traditional skills of the literary critic and then dares to join several distinct bodies of work which seldom share the same scholarly consciousness. This exhibition of intellectual cross-breeding is stimulating to watch as it discloses itself, pointing out connections and relations which specialists cannot see. Who else, after all, would think, "*McTeague,* in my opinion, provides a near perfect fictional elaboration of Catherine MacKinnon's brilliant theories on female sexuality and commodification" (p. 384, n.87); or would commit a three-page endnote (pp. 380–382) to analyzing Bryan Singer's innovative film, *The Usual Suspects,* as an example of ritual practices among youth groups in theoretical relation to Robertson Smith's 1889 lectures which Durkheim found so useful.

All this presupposes a great deal. First, one must know that Frank Norris's *McTeague* (1899), a novel of flagitious dentistry—"probably the most gruesome work in the American literary canon" (p. 83)—can be reasonably read, as Mizruchi explains, for its implicit theory of sacrifice as the necessary road to some small measures of pleasure. Then the attentive student of such syncretism must evaluate MacKinnon's notorious ideas, not only for themselves in today's heated atmosphere of feminist legal theory, but also vis-à-vis Norris's century-old novel as it relates to the ideas of social science in his own period. If one makes the time for this sort of pursuit (which in its far-flung demands only hints at what is required to follow the book sequentially), one begins to understand the great virtues to this type of scholarship as well as its obvious costs, both for the author, who committed 10 years to the book, as well as her scarce, diligent reader.

Yet the benefits of having Mizruchi's efforts available are legion, as this typical summarizing passage makes plain:

> Victor Turner has even suggested that sacrifice has "a certain generosity."
> Yet when we are made to confront the specifics of blood sacrifice, a different order of judgment is introduced. When sacrificial destruction is seen as the expression of a particular politics, in which whole groups are categorized as expendable while others are designated as beneficiaries, the more generous aspects of the rite begin to disappear. No matter how fully camouflaged it is by a bureaucratized role in a modern system of exchange, regardless of the context in which it appears, sacrificial thinking invariably reduces at some points to its fundamental identification with ritual violence. To invoke the word "sacrifice" at the turn of the century in particular was to call up a familiar but disturbing constellation of events, available in textual form through the narrative record of a biblical past and through ethnographic accounts of a primitive present. To invoke the word was to commit oneself to solemnity. In this context, sacrifice implied spirituality bordering on peril.
> (P. 30)

"Spirituality bordering on peril" is a type of rhetorical insight more readily given to literary critics than to social scientists and helps account

for the enduring legacy of criticism-qua-artistry in which Mizruchi has taken her place. Her main point on strictly theoretical grounds—that sacrifice holds societies together by victimizing one category of creatures or people to the benefit of the collective—is unvarnished Durkheimianism. But whereas it is comfortably simple to learn or teach when taken from early functionalist manifestos, its gory reality only comes home to its full extent when unorthodox works are examined that revolve around the sacrificial motif. (For the biographical background, see the feature article on her work and this book in the *Chronicle of Higher Education,* June 12, 1998, pp. A15–A16.)

For instance, when Billy Budd is shamefully hanged so that the social integrity of his ship can be maintained, Mizruchi sees it this way: "*Billy Budd* confirms not only a formal but an historical continuity: the dominance of sacrifice as a social practice. For if there is a single message in my book, it is the relevance of sacrifice as social thought and social action, supporting the most entrenched as well as innovative institutions (from charity to life insurance) and mediating the most complex developments (from the 'invention' of homosexuality to the rise of racial segregation)" (p. 23). This is excellent social theory—even if its creator is an English professor. Perhaps we should enroll in her class, reconceiving social thought along the fruitful lines she has drawn.

*Losing Legitimacy: Street Crime and the Decline of Social Institutions in America.* By Gary LaFree. Boulder, Colo.: Westview Press, 1998. Pp. xvi+240. $27.00.

James F. Short, Jr.
*Washington State University*

Documenting and describing trends in crime is difficult enough, while relating them to other social trends and interpreting the results is infinitely more so. Problems of data quality plague those who would attempt the former, and problems of inference and theory plague those who attempt the latter. Such is the case for this volume, which is the culmination of Gary LaFree's research program (with Kriss A. Drass) focusing on long-term institutional effects on crime rates.

LaFree's thesis is that the rapid increase in post–World War II street crime rates in the United States was produced by a crisis in institutional legitimacy; specifically, in the declining legitimacy of political, economic, and family institutions. LaFree carefully explains his general thesis, strengths and weaknesses of ways of measuring street crime trends in the postwar United States, and what the data suggest regarding offender characteristics during this period. His examination of common explanations of crime finds them wanting but useful within the context of the institutional changes that are at the heart of his theory. Defining institutions as "the patterned, mutually shared ways that people develop for

living together" (p. 71), in separate chapters he examines evidence of the declining legitimacy of the aforementioned institutions. The data are familiar to those who follow data on civil strife, litigation, electoral participation, polls concerning trust in institutional leaders, income and economic inequality measures, and measures of family composition and births out of wedlock.

LaFree gives special attention to the seeming paradox that crime rates continued to increase even as expenditures for criminal justice, education, and welfare increased. The answer, he suggests, is that these expenditures occurred only after crime had begun to accelerate and that their effects were muted or neutralized by *intraracial* income inequality and other events affecting this country's political, economic, and family life.

The book's uniqueness and strengths—its focus on national crime and institutional data, on comparison of time series, and its focus on accounting for African-American/white differences—foreshadow its limitations. National data tell us little about what goes on in communities (where much of the experimental and quasi-experimental "action" in crime control occurs these days) or in groups associated with subcultures. Focusing on political, economic, and family institutions—important as they are— diminishes, by implication, the role of other institutions and other social changes (e.g., voluntary associations, religious institutions, illicit as well as legitimate opportunity structures, and youth subcultures). Data on African-American/white differences, though they are the most reliable among problematic ethnic and racial categories, ignore other categories that are of increasing importance at this stage of our history as a nation.

LaFree struggles to compensate for these limitations by occasional references to other racial/ethnic categories and to studies using other than national data. Fundamentally, however, while institutional legitimacy is profoundly important to social change, in the end, the scope of the book is necessarily quite limited. Although he acknowledges the challenge to institutional controls posed by "the unique rise of a 'youth culture'" (p. 74) in postwar America, for example, the implications for crime among young people, richly documented by others, is ignored. And why did "crack" cocaine become "uncool"? (This transformation is, perhaps, a major reason for declining homicide rates among young African-American men.) Similarly, global political, economic, and social changes that have an impact on institutional legitimacy are given short shrift.

The book seems clearly to be intended for student consumption and, perhaps for this reason, the presentation often seems simplistic. A more nuanced consideration of institutional legitimacy might, in addition to considering a broader range of institutions, distinguish between the significance for criminal and other types of behavior, of varying meanings of trust (competence, fiduciary responsibility, and the integrity of the moral order), and their relationship to institutional legitimacy, as Bernard Barber has argued (*The Logic and Limits of Trust* [Rutgers University Press, 1983]). Still, there is much of value in exposing younger generations, unable to be living witnesses, to the postwar period, the turbulent '60s, and

the beginnings of the global political, economic, and social developments that now engulf us. To his credit, LaFree has been persistent in confronting truly important questions, surely a quality to be admired.

*Shifting the Blame: How Victimization Became a Criminal Defense.* By Saundra Davis Westervelt. New Branswick, N.J.: Rutgers University Press, 1998. Pp. xi+193. $52.00 (cloth); $22.00 (paper).

Ross Macmillan
*University of Minnesota*

The repeated (and repeated and repeated) images of Lyle and Eric Menendez crying on the witness stand as they recounted tales of emotional, physical, and sexual abuse at the hands of their parents were profoundly unsettling. While these images and stories were themselves disturbing enough, social critics and legal scholars expressed additional concern that we were witnessing a new and potentially destructive defense strategy, a strategy that blurred the line between victims and offenders and excused violent, criminal behavior. This, the emergence and use of the "victimization defense" or the "abuse excuse" in American courts, is the subject of *Shifting the Blame.*

Westervelt's approach to examining the victimization defense is two-pronged. First, she examines the implications of victimization defenses for contemporary legal discourse, particularly notions of legal responsibility. Second, she examines appellate cases to map out the emergence of, situational character of, and successes and failures of victimization defenses. Ultimately, she seeks to develop "a sociological framework for explaining a specific type of legal change—the development of the victimization defense strategy—and use that change to say something about the larger culture in which we live" (p. 12).

Well organized, each chapter of *Shifting the Blame* is framed around a specific and distinct research question. After an initial introduction, the second chapter begins by examining the conditions necessary for the attribution of criminal responsibility, describing different defense strategies for challenging legal responsibility (i.e., self-defense, insanity), and suggesting connections to victimization defenses. The third chapter examines the origins of "victim-based" strategies, which were initially deployed in cases involving economically deprived defendants, victimization by a "rotten social background," those coercively persuaded, victimization through "brainwashing," and victims of spousal violence. Westervelt points out that the victimization was really only successful in the latter case; efforts to consider social victimization and coercion as victimization defenses generally met with little success.

Attempting to explain these differences is the objective of the fourth chapter, and it is here where an attempt is made to more explicitly examine the relationship between law and society. Specifically, Westervelt ar-

gues that three extralegal factors, "reformers, resources, and social context" (p. 82) account for the success of the victimization defense for battered women. Feminist reformers, lawyers, and social scientists who supported the legal defenses of battered women were able to mobilize organizational resources. Furthermore, these advocates were able to link fundamental themes of the defense to broader and more familiar concerns of gender inequality. Herein, argues Westervelt, lies the explanation for the success of the victimization defense for battered women.

While this approach provides a potentially useful framework for understanding the sources and mechanisms of legal reform, her analysis is ultimately a bit sketchy. Within the context of a 31-page chapter on the role of reformers, resources, and social context in the successful recognition of the victimization defense for battered women, a scant page and a half are devoted to comparative analysis with economical deprivation and coercion cases. It is argued that defense strategies in the latter cases simply failed because they were unable to draw organizational support and unable to link their defense strategies to general cultural concerns. This lack of detailed consideration ultimately weakens Westervelt's theoretical agenda. There simply is little consideration of the role of social context and resource mobilization in the *failure* of the other types of victimization defenses. Yet, social context and cultural conditions should both facilitate and impede processes of legal reform.

Consider, for example, cases based on claims of economic deprivation or "rotten social background." The emergence of these cases coincided with quite dramatic changes in public and political discourse on social and racial inequality. Johnson's War on Poverty gave way to Nixon's War on Drugs, which then gave way to the War against the Poor (Herbert Gans in *The War against the Poor: The Underclass and Antipoverty Policy* [New York: Basic Books, 1995]). Examining how social context and public discourse may have contributed to the failure of the victimization defense for such cases would have both strengthened and broadened Westervelt's analysis of the social sources of legal reform. Ultimately, this would have presented a stronger theoretical argument about the interconnections of law and society.

After describing the expansion of the use of the victimization defense to include new types of defendants, new types of cases, and new forms of victimization, the book concludes with further consideration of connections between law reform and society. Westervelt suggests that the legal trials over the adoption of the victimization defense are reflections of a "culture struggling with a new identity—the identity of 'victim'" (p. 143). In victimization defense cases, law and society interacted in complex ways as both struggled to integrate old notions of responsibility with new notions of victimhood. The story of the victimization defense is ultimately a story about redefinition within law and within society.

As a study of legal reform, Westervelt's research provides a unique and interesting look at legal change, particularly the ways in which competing definitions of victims and offenders challenge and are challenged by clas-

sic notions of legal responsibility. As research on the links between society and law, it falls short in fully articulating the extralegal and societal influences on legal reform. Still, *Shifting the Blame* is an interesting and topical analysis that extends previous understandings of the social determinants of legal reform.

*The Public Realm: Exploring the City's Quintessential Social Territory.* By Lyn H. Lofland. New York: Aldine de Gruyter, 1998. Pp. xix+305. $52.95 (cloth); $25.95 (paper).

Mark Gottdiener
*University at Buffalo*

Poorly edited, undertheorized, and self-indulgent to the extreme, Lofland's *The Public Realm* is less a contribution to urban and cultural studies than a cautionary tale. On page 144 she writes: "Think back to Chapter 1. . . . As I said there," which is followed by an extended, single-spaced quote of material equivalent to an entire page of text. Sure enough, I found the same sentences on page 8. While many of us reluctantly request the reader to recall previously discussed material, never before have I encountered an author who quotes at length from an earlier page. Does the publisher not employ editors? Are these ideas so profound that they require self-quoting? Hardly. This book is an amplification of a 1970s work that focused on observations of people's interactions in a city environment. The people are strangers and pedestrians. Their interaction in urban public space is certainly very interesting. But Lofland asserts that this kind of microstudy of interaction without social contextualization of any kind is, in effect, the *most* important and central study of the city. From this starting point she not only dismisses well-known urbanists, such as David Harvey, using simplistic ideological labels, but she ignores the relevance to the study of public space of culture, class, race, gender, age, and ethnicity, not to mention larger social forces such as juridicical, political, and property relations that regulate public space.

This book has an overabundance of references, but it is highly unlikely that Lofland thinks very much about this material. How else to explain the listing of Habermas's seminal work on the public sphere in the bibliography, but the complete absence of any theoretically informed analysis of the "public sphere" concept in contemporary society; nor does she discuss the changes that have taken place with the advent of modernity in legal and property relations. In short, this book is not about the "public realm" in any kind of sophisticated, theorized sense, despite the title. She explains virtually all the socio-spatial changes brought about by the industrial revolution, such as the separation of home from work, simply by a descriptive dichotomy she calls enlargement/enclosure. No need, for example, to discuss the role of economic, political, power, and cultural factors in this same historical transformation. Counter examples to this

observation, such as the case of Beijing in the 1500s (which was both enlarged and enclosed yet without an industrial revolution), also do not seem to be relevant.

In the earliest chapters she defines the object of her study, namely, the public realm, unlike Habermas, simply as an aspect of the environment, that is, without recourse to any juridicopolitical distinction or its changing historical significance. What she then means by calling chapter 2, "the normative or 'legal' system" I do not know, because the legal aspects of public space are simply disposed of in a few paragraphs. She quotes Elias but seems not to understand the significance of his work that exemplifies the articulation of the macro with the micro, of state power with public etiquette, for example. We could have benefited greatly by a historical discussion of the changing mode of public encounters across history and their mechanisms of regulation, topics that might be covered by a deeper book on the public realm.

Chapter 3 reprises Lofland's earlier analysis on the systematic observations of strangers. There is interesting material referred to here that has been produced by symbolic interactions, such as Fred Davis and Lyman and Scott, over the years (see, e.g., Fred Davis, "The Cabdriver and His Fare" *American Journal of Sociology* 65 [1959]:158–65 and Stanford Lyman and Marvin Scott "Territoriality A Neglected Sociological Dimension" *Social Problems* 15[1967]: 236–49). The remaining chapters are sheer self-indulgence. Without establishing a deep level analytical foundation, Lofland offers us her opinions on "esthetics," specifically, those aspects of the built environment that are "enjoyable," a purely subjective discussion. The final chapter, "The Uses of the Public Realm" raises an interesting, even critical issue, namely, the problematic character of public space and the shrinking of the public realm, especially for U.S. cities. But, this discussion is cast purely in physicalist terms. Both social relations and processes are ignored. The author raises the important topic of how urban planning and design can be improved by greater attention to the public realm. However, without mention of the larger social forces conditioning space and the relationship between the macro and the micro aspects of the street, this final, concluding chapter falls short of substance.

*Reconstructing Chinatown: Ethnic Enclave, Global Change.* By Jan Lin. Minneapolis: University of Minnesota Press, 1998. Pp. xix+248. $19.95 (paper).

Pyong Gap Min
*Queens College*

Chinese communities in American cities are unique compared to other ethnic communities because they have their own enclaves, commonly referred to as "Chinatowns," established usually inside central cities. More than a dozen books on Chinatown have been published in the last two

decades, and each one has a different foci and a different theoretical orientation. For example, Min Zhou's *Chinatown* (1992) emphasizes the socioeconomic potential of New York's Chinatown using the enclave economy thesis. In stark contrast, Peter Kwong's *New Chinatown* (1987) revolves around class conflicts and labor movements within the same community.

Based on ethnographic research and public documents, Jan Lin's *Reconstructing Chinatown* covers most of the topics covered by Kwong, for example, the labor movement organized by Chinatown workers, the inflow of overseas Chinese capital into Chinatown, and community activism. However, it differs from *New Chinatown* and other books on New York's Chinatown because of Lin's "global city" perspective in analyzing the enclave's economic, social, political, and cultural life.

Lin devotes a significant portion of the book to examining the impact of the inflow of Chinese capital and labor from Hong Kong, Fujianese ports, and other Southeast Asian countries into New York's Chinatown on community changes. Lin uses his "global city" perspective most effectively in chapter 3 when he examines the impact of transnational (Asian overseas Chinese) capital vs. local capital on banking and real estate industries in Chinatown. Lin documents in greater detail the investment of overseas Chinese capital in Chinatown businesses in banking, real estate, and other industries, and the increasing globalization of the Chinatown banking industry, using both qualitative data and statistical figures. His historical overview of the formation of overseas Chinese capital is especially informative.

The Chinese immigrant group is more socioeconomically polarized than any other Asian immigrant group; meanwhile, the Chinese community in New York City has a higher level of class polarization than any other Chinese community in the United States. Class conflicts among Chinese immigrants in New York City are manifested most visibly in the labor struggles of Chinatown sweatshop workers, toiling in Chinese restaurants and garment factories. Lin examines in considerable detail labor struggles among garment and restaurant workers in Chinatown from the early 1980s to 1995. He also analyzes a 1995 demonstration among Chinese garment workers in Sunset Park, a satellite Chinatown in Brooklyn. By revealing the active participation of Chinese workers in the labor movement, Lin seeks to dispel the popular beliefs that new immigrant and illegal resident workers are generally docile and that the massive immigration of people from less developed countries into the United States hurts the interests of the American labor movement (p. 192).

Many Americans perceive Chinatowns as ethnic enclaves isolated from the rest of society. However, Lin's analysis of various cases shows that Chinatown residents, merchants, and officials interact actively with local, state, and federal government organizations. For example, as a result of rezoning recommendations by the New York City Department of City

Planning, several high-rise buildings have been constructed in Chinatown since the mid-1980s (pp. 151–56). Chinatown residents initially challenged in the courts the city's plan to construct high-rise condominium projects on the grounds of tenant harassment. This opposition is one of the examples of the community's conflicts with the local government.

Chinatown has long been perceived as "a district pervaded with organized crime, vice industries, and depravities associated with illegal immigrant smuggling and sweatshop activity" (p. 171). Lin points out that these images have been reinforced by newspaper tabloids, police and detective television series, and Hollywood films made to meet the "voyeuristic impulse" of white Americans. In chapter 7, he provides a review of major musical shows, books, and Hollywood films that stereotyped life in New York's Chinatown and Chinese American characters since the late 1880s. He introduces the periodical protests of Chinese Americans and the Chinese government against these cultural stereotypes. Lin also covers the recent efforts of new Chinese American filmmakers to deconstruct popular stereotypes. In sharp contrast to the stereotypical accounts he cites, Lin presents life in New York's Chinatown as complex and vital, and as a researcher, he joins the efforts to deconstruct popular stereotypes of Chinatown and Chinese Americans.

Lin's book has minor problems in terms of data and organization. For example, he indicates that "there were about 120 Chinese stores and businesses located in Flushing in 1982 in a diverse range of industries" (p. 116). For a book to be published in 1998, he should have updated the data in this instance as well as in a few other places. As of March 1999, I enumerated more than 200 Chinese businesses in Flushing. Chapter 2 focuses on labor struggles among Chinese sweatshop workers in restaurants and garment factories. He includes in the chapter a section on Chinatown street traders, despite his acknowledgement that, as self-employed operators, street traders have no history of labor-capital conflict (p. 76). The data on the street peddlers might have been more appropriate in chapter 1, in connection with business profiles of Chinatown, or in chapter 6, in relation to the street peddlers' confrontations with the city police and regulatory agencies.

Despite these minor limitations, *Reconstructing Chinatown* is a significant work that is relevant in several fields. First, it is probably the best book on Chinatown because it covers several critically important topics in examining life in Chinatown, utilizing both quantitative and qualitative data. Second, it is a great contribution to the race and ethnic relations field because its deconstruction of negative stereotypical images provides an alternative model for examining immigrants and ethnic enclaves. Finally, this book has nicely connected the field of race and ethnic relations with urban studies by examining an ethnic enclave with the "global city" perspective. This book is ideal as a textbook for courses related to Asian American communities, urban studies, and race and ethnic relations, particularly at the graduate level.

*King Kong on 4th Street: Families and the Violence of Poverty on the Lower East Side.* By Jagna Wojcicka Sharff. Boulder, Colo.: Westview Press, 1998. Pp. x+258. $65.00 (cloth); $18.00 (paper).

Ida Susser
*Hunter College*

Since New York City announced bankruptcy in 1975, the historic commitment of its administration to public education, public health, and social services has been replaced by an effort to remove the poor from potentially high-rent, upscale areas. With a few exceptions, such as Central Harlem, Manhattan could on this trajectory become a bastion of privileged residency. As the number of children living below the poverty line has risen and public assistance entitlements have been reduced or cut off, even the poverty-stricken Lower East Side has become "revitalized" with rising rents and real estate values.

Jagna Sharff's insightful and detailed ethnography of this area spans the process of the neighborhood's transformation from a historic tenement neighborhood of immigrants and working poor to the gentrified home of well-to-do artists and the hip elite. The book dramatically captures the impact of this transformation on the lives of poor New Yorkers.

Sharff began her research in the 1970s, and much of the power of her account derives from the fact that she has followed many of her informants as they struggled through the next 20 years. Sharff established an office in the first floor of a rundown apartment building on the Lower East Side and created a center where a team of researchers could meet among themselves, interview people, and "hang out" with local residents. First to take advantage of this new resource were the neighborhood children, and the researchers swiftly recognized them as a key entry point into the complex life of the community. Through descriptions of the children as they grow from charming and mischievous youth through troubled adolescence to the prison experiences of their adulthood, Sharff achieves an account that vividly documents institutional barriers and the destruction of the lives of hopeful young men. One of the most striking narratives in the book concerns Miguel who is named "the most improved student" and must give a speech at graduation from his sixth grade public school class. Miguel strives to succeed as a dutiful son while his brothers resort to selling drugs accompanied by the inevitable prison terms. Eventually, in response to personal alienation and poverty, Miguel too turns to the drug trade and soon finds himself in solitary confinement in one of the most dangerous prisons in New York State.

Sharff documents the resilience of mothers as they watch the street, learn from other women, and try to survive in a city where landlords welcome drug dealers for the higher rents that they pay but mothers have to cope with the violence and recruiting of their children that the dealers bring. Beyond the poor themselves, we meet the landlords, the local activists, and the *bodega* owners and, through Sharff's skillful narration,

watch larger socioecanomic forces—how the extensive arson of the 1970s and the changing real estate patterns have worked themselves out in the deaths of children and the destruction of a vibrant poor community.

Sharff is careful to describe the varied backgrounds of her researchers, who almost become characters in the book. She talks about the methodological discussions and doubts that her research team confronted. Would interviewing teenagers and neighbors immediately after the unexplained murder of a young man on the same block endanger their informants or themselves? How can they help the local children through a long hot summer when school supervision and lunches are not available? This detailed focus on people who came to the office or lived on the block has its disadvantages. It leads to some gaps in the analysis of wider issues such as city policies or the perspectives of residents beyond the Puerto Rican population. But Sharff is aware of the limitations of her method, noting her ignorance about the source of many disputes, including the context of a child's murder in the corner grocery store. The brothers who own the store, Dominican, as opposed to Puerto Rican immigrants, seem to be outside the team's communication network. Indeed, it is partly Sharff's awareness of the partial knowledge the research team has gathered that imparts integrity to this account.

The book documents the lives of victims of changing city, state, and federal policies without either "blaming the victim" or denying the agency of the sufferers. Sharff's descriptions of women struggling to find jobs, better housing, and affordable nutritious food and to keep their children in school forcefully undercut stereotyped versions of "the underclass."

Overall, this rich ethnography allows for interpretation of many issues beyond the changing urban neighborhood. It contains within it a sensitive rendition of the complexity of women's lives particularly as they confront the continuing drama of mothers and sons. *King Kong on 4th Street* captures New York City at a transformative moment and portrays the hope and the tragedy of the families who lived on the Lower East Side in the 1970s and 1980s.

*The Unknown City: Lives of Poor and Working-Class Young Adults.* By Michelle Fine and Lois Weis. Boston: Beacon Press, 1998. Pp. x+328. $24.00.

Paul A. Jargowsky
*University of Texas at Dallas*

In *The Unknown City,* Michelle Fine and Lois Weis attempt to paint a realistic portrait of life in urban America, based on extensive life history interviews and focus group meetings with members of Generation X who are rarely heard from: the poor and working class. The authors' goal is laudable, given the extent to which so much urban research continues to rehash the same lifeless census data or panel surveys. Research based on

these data sets, as important and necessary as it is, can never capture the qualitative experience of those who struggle to survive in high-poverty neighborhoods. Unfortunately, the book suffers from a number of serious flaws. First, Fine and Weis employ a survey instrument that seems designed to encourage their subjects to speculate about general theories of the economy and society rather than to tell us about the concrete conditions of their own lives. Second, the voices of the urban poor are all too often drowned out by the didactic lecturing of the authors. Despite these flaws, the book may still be useful to urban scholars who are willing to read around the Fine and Weis text and draw what insights they can from the voluminous quotes of their research subjects.

Fine and Weis set out to create "a biography of urban America" (p. 1). They conducted in-depth interviews, both individually and in focus group settings, with 154 young adults in Buffalo, New York, and Jersey City, New Jersey. The subjects are not meant to be representative of their cohort but rather "young adults who basically 'followed the rules.' " Moreover, all interview subjects were involved in community institutions, such as schools, churches, agencies, and community centers, through which the sample was obtained. The interview pool consisted of persons suggested by "gatekeepers," that is, staff of the community institutions, and was expanded later through snowball sampling. While such selection could well be justified depending on the aims of the research, Fine and Weis downplay the selection issue and treat their subjects as generally representative of the young urban poor.

The interview protocol is particularly weak. The questions encourage interview subjects to speculate at a high level of generality. For example, there are two questions on gender. One asks subjects to talk about "the images of men and women in this community." The second asks subjects to "describe relations between women and men" in their community (p. 302). In actual practice, the interviewers went even further down this road; for example, one questioner asks a female subject, "Why do you think some men are beating up women? We hear a lot about this" (p. 126). Speaking in such general terms will tend to elicit the subject's preconceptions, shallow generalities, and the "conventional wisdom" rather than the raw material from which a social scientist could construct a nuanced view of gender relations in poor communities.

Fine and Weis succeed in letting us hear the voices of the urban poor by presenting numerous extended quotes from their interviews. A series of chapters cover the state of the economy, race relations, sex and sexuality, and crime and violence. Somewhat reluctantly, given their view that race and gender are socially constructed categories, they present the views and experiences of the subjects ordered by race/ethnicity and gender categories. This structure inevitably leads them to make sweeping generalizations along racial and gender lines. The following is typical: "Rather than offering a critique of white elites who have directed the restructuring of the American economy, rendering the white working class superfluous, the white male version of social critique is, by and

large, a clear, targeted, and critical litany of judgement on the focused actions/behaviors of others, particularly African Americans, and particularly men" (p. 21). White men are described as "filled disproportionately with discursive venom that is difficult to hear" (p. 131). In contrast, several pages later black males are lauded as providing a "broad based social critique . . . of racism and society" (p. 28). Even the most banal comments are depicted as insightful social and economic "critiques" of an oppressive system of white privilege and economic domination. In service of this political agenda, the authors seem to gloss over or miss entirely many of the more subtle and interesting aspects of the interview fragments presented.

In the chapter regarding white women's experience of domestic violence, Fine and Weis categorize their respondents into "hard living" and "settled lives" women, differentiated by the stability of their family structures and sources of income. The former group readily reports domestic violence, both in childhood at the hands of their fathers and in adulthood from husbands and companions. However, by and large, white women rarely "voice some critique of men and family" and express a "tone of reconciled contentment in which they wrapped narratives of the family as loving and supporting" (p. 147). Fine and Weis assume these women are covering up domestic violence: "Domestic violence is a domain of muted critique that has difficulty breaking through" (p. 155). They argue that the silence about domestic violence from the women leading settled lives reflects a way of "valorizing" the family as a marker of whiteness that differentiates them from minority groups. Perhaps—but it could also be true that white women living settled lives actually experience lower rates of domestic violence than hard-living women. How Fine and Weis turn the absence of evidence into evidence is troubling and suggests an unwillingness to hear those parts of the subject's answers that conflict with their preconceived ideas about the subjects' lives.

*The Unknown City* represents an outstanding idea—to understand the world of the urban poor as they themselves perceive it—that is poorly executed. The voices do come through, and that is the book's great contribution. Unfortunately, these voices must compete with Fine and Weis themselves, who continually use and sometimes misuse the respondent's words to support a particular critique of the U.S. economy and social policy. Even for those of us who agree with large parts of this critique, the constant repetition seems tiresome and often forced. The book could have and should have told us more about the heterogenous experiences of the urban poor and less about the political ideology of Fine and Weis.

*Don't Call Us Out of Name: The Untold Lives of Women and Girls in Poor America.* By Lisa Dodson. Boston: Beacon Press, 1998. Pp. x+256. $24.00.

Ruth Horowitz
*New York University*

This rich set of interviews of 50 poor women, discussions from focus groups, and responses to surveys give the reader an ample picture of the lives of women who are trying to survive and raise children alone in today's service economy. Few jobs pay well enough to take these women out of poverty, but many of them persevere, gain an education, and find jobs that allow them to support their families and give them a stronger sense of self. While most of the women with whom Dodson had many conversation over several years are not the poorest or the least educated, these are women and their children who have to use welfare and government services to survive, sometimes more than once, as they work hard to establish themselves in a society that typically fails to respect them.

While each woman's experience had unique elements, one of the most interesting themes developed concerned the women's experiences growing up typically in strong family networks, gaining their education, and becoming mothers themselves. As they were growing up, many of the women interviewed had helped their mothers to raise the younger children and to do household chores. There were no babysitters, au pairs, or maids to help with the daily needs of a growing family with working parents; the daughters took that responsibility. Although these responsibilities provided them with a sense of their own capabilities, they did not have the luxury of a prolonged childhood. Some resented the significant time spent doing household tasks and child care, and some did not have the opportunity to invest sufficient time in their schooling. Mothers generally supported their daughters' education when they were younger and were proud of their achievements, but it was often family responsibilities that resulted in their dropping out. Moreover, when some daughters wanted to continue their education beyond high school, it was not understood, and they received little support from family or boyfriends.

When they are expected to and do hold adult responsibilities—they are very self-reliant—they want to be treated as adults. Moreover, so embedded in family relationships and responsibilities in caring for others, it becomes difficult to imagine oneself in a different sort of role. In fact, the young women imagine themselves as mothers, and those who take jobs or work toward them chose careers and work related to the care of others. Many, however, have a child before they obtain a significant work history. It is, from their experiences, the most important responsibility for a woman, and many cannot really imagine another life.

Despite hopes for ongoing relationships with the fathers, few have them, and after they have the child, many feel isolated. This loneliness appears largely as a transitory stage, unlike some of the other recent re-

search on young mothers that provides an image of women suffering from lack of support and strong relationships growing up, especially after the birth of the baby. When their babies are very young, many of the women interviewed by Dodson used welfare and other state benefits to support themselves and experienced depression and the degradation that goes along with welfare use in the country, though many have had good experiences with some thoughtful social service providers, teachers, or counselors. Welfare is not a voluntary resolution, however. It too, for most of the sample, is a transitory part of their lives, though some do return to using it more than once. They do what it takes to survive. According to Dodson, most had very significant attachments when younger, and as they became more settled and struggled to make a place for themselves in the world of work and motherhood and managed the delicate, difficult, ever-shifting balance between the two, they developed new relationships and reaffirmed old ones.

This does not mean that all their relationships are positive. Dodson deals directly with and demonstrates the complexities of many of their relationships of abuse both by men in their communities and by men employed in the welfare and social service systems. Many were subject to considerable abuse both as very young women and later as adults. Some are too tired to protect themselves and their children from abuse. Others submitted to demands from service providers because they needed the services, while others suffered because they refused to accede to demands. Dealing with relationships with men is complicated and rarely smooth.

The purpose of this book is to present readers with variations in the serious struggles of women to make it when they are largely alone and without access to jobs that pay a reasonable wage to support a family. It shows the inconsistencies of a welfare system that forces women to go back on welfare when their child care benefits end and shows how they must find innovative strategies to manage with children and work. Readers who are interested in women's stories will be fascinated by the richly textured tales of struggles and some triumphs over the odds. But others who are interested in understanding, for example, the kinds of experiences that permit some to triumph while others fail will find the structure of the book a bit frustrating. The range of solutions to problems in and interpretations of family, motherhood, sex, work, and friends are not always transparent, though it is clear that they are not the same for each woman. In her efforts to see the unique in everyone, there is a limited presentation of that which can be compared and generalized. Stereotypes of women living in poverty are destroyed, and a richly textured account of many women who have lived much of their lives in poverty is constructed. This is the success of this book.

*From Stumbling Blocks to Stepping Stones: The Life Experiences of Fifty Professional African American Women.* By Kathleen F. Slevin and C. Ray Wingrove. New York: New York University Press, 1998. Pp. viii+187. $50.00 (cloth); $17.50 (paper).

Eleanor Palo Stoller
*Case Western Reserve University*

Kathleen Slevin and C. Ray Wingrove illuminate the complex interplay of history, systems of inequality, and individual biography in their compelling qualitative study of 50 retired professional African-American women. The title, *From Stumbling Blocks to Stepping Stones,* reflects the themes of survival and resistance through which these women interpreted their biographies. The authors use a life course framework to interpret the qualitative data generated through life history interviews.

A major strength is the weaving of narratives to capture the influence of family, school, church, and community on the remembered worlds of these women's childhoods. Memories of childhood include lessons of survival and resistance taught by fathers as well as mothers. Echoing a prominent theme in the literature on families of older African-American women, these women describe parental efforts to create both emotional sanctuaries and armors or resistance. Lessons learned in the home were reinforced in the churches that were central to their social lives as children. It was in the church that they learned etiquette, honed their leadership skills, and gained self-confidence. School teachers supplemented the regular curriculum with black history and instilled "pride in self and race." While describing the ethos of sharing that characterized the segregated communities in which these women grew up—an ethos that was "both a form of resistance and a celebration of humanity at its best" (p. 75)—Slevin and Wingrove document the ramifications of intersecting hierarchies based on race, gender, and social class. Awareness of gender inequities is a product of hindsight, largely invisible during their youth, when the ubiquitous impact of race dominated their awareness. Indeed, the women who became most acutely aware of gender discrimination were those who worked in predominantly African-American settings, a finding that reinforces the importance of considering context in understanding the dynamics of inequality.

Little of this is new. Researchers have demonstrated that the church was the black community's chief institution of power, that families provided children with emotional sanctuaries from racism, and that the black communities reinforced an ethos of "giving back" and "lifting while you climb." What Slevin and Wingrove add to this established history are compelling insights into the lived experiences of women whose lives crosscut periods of dramatic change in the shape and dynamics of inequality within the United States. By giving voice to women whom they describe as "pioneers in professional fields at a time when racial and gender barriers were pervasive" (p. 161), the authors undermine stereotypical

images of elderly black women as poor, sick, and uneducated. Without romanticizing the barriers imposed by hierarchies of race, class, or gender, the authors document the childhood lessons, community support, and adaptive resources that enabled these women to craft retirement years characterized by financial security, pride in their achievements, and a sense of obligation to "pass on their legacies of survival and resistance" (p. 123) to the black community.

Slevin and Wingrove are less convincing in their claim that these women can serve as role models for future cohorts of women, who will also combine child rearing with continuous labor force participation. Except for a reference to "sojourner mothers" who left their spouses, infants, and children to finish degrees, there is no discussion of how these women balanced the competing demands of paid work and family responsibilities. Neither do they address the sociohistorical context in which different cohorts of women experience particular stages of the life course.

Although guided by a life course perspective, the authors argue that the "pervasive and homogeneous" impact of "racism, intertwined with systems of class and gender inequality" negated "the usefulness of cohort analysis" (p. 8). Without more evidence, it is difficult to accept the claim that historical context had no differential impact on the biographies of women born from 1907 to 1941. Were the work experiences of the women who entered the labor force during the depression of the 1930s similar to those who began their careers in the expansionary economy of the 1950s? How did the demands for labor during World War II affect employment opportunities of the women entering the labor market during the 1940s? The authors argue that opportunities opened up in the 1960s for some African-American women who worked in predominantly white settings. Is it true that these changes had the same impact on the career trajectories of women in their 50s as on women in their 20s and 30s?

Despite these unanswered questions, Slevin and Wingrove craft a richly textured collective biography illustrating both survival and resistance. Evidence of inequality from childhood into retirement reminds readers that racism is a legacy that continues to shape the lives of these women. Personal narratives of the historical time captured by these lives offer nuanced accounts of the progress Toni Morrison characterized as the "slow walk of trees," as well as feelings of loss of the shared mission that characterized the segregated communities of their childhoods.

*Double Burden: Black Women and Everyday Racism.* By Yanick St. Jean and Joe R. Feagin. Armonk, N.Y.: M. E. Sharpe, 1997. Pp. xiii+235. $35.00.

Lynda F. Dickson
*University of Colorado at Colorado Springs*

*Double Burden,* by Yanick St. Jean and Joe R. Feagin, contributes to the growing body of literature on the continuing significance of race for middle-class African-Americans. Unlike other works, this book strives to place *women* at the center of the analysis and to document the continued discrimination, misconceptions, and distortions they face in contemporary society. The authors argue that while these women appear to have "made it," they routinely receive assaults at work, through negative media images, and in other areas of their lives.

The authors draw from in-depth interviews from a 1988–90 national study of more than 200, mostly middle-class African-Americans, as well as more recent focus group interviews with African-American professionals. From these sources, we learn of the discrimination and stigmatization of black women as well as the various strategies employed to "fight back."

St. Jean and Feagin begin by focusing on the work setting, specifically the various manifestations of "gendered racism." The assumption by white workers that black women are "two-fer's" or "affirmative action hires" means that their competence is always in question. Black respondents describe the "surprise" by whites when they are competent, the constant feeling of being "tested" for competency, or being made to feel stupid. Strikingly, the bulk of the chapter focuses on the perception of the (mostly male) respondents that black women are preferred over black men in professional settings. Taking this perception as a "given," with virtually no discussion of the existing literature on the subject, the authors proceed to offer two reasons why this might be the case: black women are more easily managed, controlled, and manipulated by white men, and black men are feared by whites. This (assumed) preference for black women denigrates black men and in fact illustrates the way in which "gendered racism" is manifested toward black men.

The authors also examine the impact of broader cultural standards of beauty on black women's self-image. Because black women are antithetical to beauty in this culture, they are stigmatized through doll play, magazine images, film, and television. Paradoxically, white men are attracted to them, albeit in a secretive, exploitative way. Following this theme, the authors identify the common myths and media images of black women— especially the sexually promiscuous "Jezebel" and the angry, emasculating "Sapphire"—and discuss the historical origins of these racialized images. According to the authors media images include the portrayal of black women as lower class (or if middle class as "superwomen"), immoral, or fat. Yet the authors present virtually no evidence of these images in contemporary media.

Relationships between black and white women are also discussed, specifically the ways in which white women discriminate against black women. Their actions may be subtle or overt, ranging from the marginalization of black women at social gatherings, or addressing black women so oddly or curtly that it humiliates, to withholding professional guidance and other employment opportunities from black women.

Throughout the above discussions it is argued that white treatment and white-controlled media images of blacks stem from a white need to reinforce the "collective memory" about the inferiority of blacks. The authors also include the ways in which black women (and men) draw from their "collective memory" to counter the above described experiences. The conceptual importance of collective memory requires that the authors discuss black families and motherhood. It is through black extended families that collective memories, including an "oppositional culture," are instilled and perpetuated among African-Americans. Mothers, grandmothers, and other women play an important role in transmitting this information. Because of the importance of collective memory, the authors argue that the small, nuclear family may not be the best family model for (even) middle-class blacks. They believe that African-Americans must maintain their historical reliance on the extended family for the provision of social supports and information needed to function in a racist society.

There are several problems with this book. First, glaring in its absence is the presentation of bodies of literature on the changing nature of racism, white attitudes toward blacks, the experiences of middle-class black families, or media portrayals of blacks. This literature in some cases refutes, but would often strengthen, the major arguments presented. Second, analyses often go far beyond, or in some cases are refuted by, the presented "data." Third, the authors routinely ignore the intersections of race, class, and gender. For example, the presumed preference for black women over men is based on a comparison of the employment opportunities of black middle-class women with those of *all* black men, rather than with black men of similar educational levels. Fourth, the book consistently fails to acknowledge that racism, media images, relationships in families, and "collective memories" are *dynamic* features in society. This failure undermines the central argument of the book, which is that prior survival strategies (oppositional culture) are still essential to the middle class. Yet, clearly there are differences between the "gendered racism" experienced by grandmothers and that of their granddaughters who are professional women today. These differences require very different "fighting back" strategies.

Finally, *Double Burden* is not really about black *women* but about black *men* and women, with women's experiences often taking a back seat. After the preface, there is no pretense that women's experiences and perceptions are central. The chapters on work and beauty draw almost exclusively on the "voices" of black men. Those who hope to see black women placed at the center of this analysis will be disappointed. In spite of these problems, St. Jean and Feagin pose many compelling questions

that deserve more careful exploration, and for that reason alone this book deserves careful reading.

*Fighting Words: Black Women and the Search for Justice.* By Patricia Hill Collins. Minneapolis: University of Minnesota Press, 1998. Pp. xxiii+312. $18.95 (cloth).

*Long Slow Burn: Sexuality and Social Science.* By Kath Weston. New York: Routledge, 1998. Pp. 255. $18.95 (paper).

Patricia Ticineto Clough
*City University of New York*

*Fighting Words* and *Long Slow Burn* return to the authors' earlier works—Collins's on black feminist thought and Weston's on gay, lesbian, and bisexual families—with the aim of evaluating the cultural criticism of the past 20 years and its effect on the social sciences—sociology and anthropology especially.

Collins begins with an account—no longer surprising but still disturbing—of the exclusion of black women from sociology. She argues that sociology defined itself in terms of this exclusion by figuring the black woman as other. As such, black women have been identified as objects of sociological research. Or, at most, they have been identified with the soft, qualitative interpretive methods marginal to sociology, when its authority and power as a discipline are grounded in hard, quantitative approaches. Collins's account of black women's past relationship to sociology is especially forceful in the way it echoes her treatment of the normative configuration of family and national ideologies, the state and civil society, and the private and public spheres, presumed in modern western sociology.

Since this configuration does not fit black women's experiences, Collins concludes that sociology has been complicitous in making the particularity of black women's exclusion from the public sphere incomprehensible and the gendered nature of black women's support of a black civil society invisible; also ignored is the surveillance of black women, which, along with the history of black women's patterns of work, have denied them the privacy of the domestic enclosure. Black feminist thought, therefore, was conceived as a departure from sociology; such a departure was necessary, at least at first, in order to enable black women to break silence and to be able to speak self-reflexively about race, gender, sexuality, ethnicity, and class. But, Collins argues that the political gains of black women's "talking back" now must be measured against changes brought on by the postmodern criticism of social science and Afrocentricism.

Unfortunately, Collins draws only on secondary accounts of the postmodern criticism of social science, so that her treatment of it is an overly familiar one. Demanding further debate, however, is Collins's assertion

that the postmodern treatment of difference is merely an individualistic approach to identity, which, in emphasizing individual choice, ignores group formation and institutional arrangements. Turning to Afrocentricism, Collins argues that in its excesses, Afrocentricism is not politically enabling for black women. She is especially critical of the implications of Afrocentric thought for black women's sexuality, their role as mothers, and their participation in civil society generally. She concludes that family ideology often functions along with nationalistic discourse to situate subjects in hierarchical arrangements of race, gender, ethnicity, sexuality, and class; family ideology also is deployed to naturalize these hierarchical arrangements.

Collins instead calls for a critical social theory that not only encourages black women's full participation in social life but actually grounds its own authority in its ability to enable black women's full participation. Collins closes *Fighting Words* drawing on the experiences of Sojourner Truth in order to rethink the connection between black feminist thought and black women's spirituality, their religious faith, and their deep commitment to the black church. Collins's treatment of Sojourner Truth will surely draw debate, and such debate will be worthwhile for addressing the question of whether a moral authority, or what might be better termed an ethics, is necessary for a critical social theory that grounds its authority in its potential to enable a just social life.

In *Long Slow Burn*, Weston also revisits her earlier work and offers an evaluation of cultural criticism of the last 20 years in relationship to anthropology. More specifically, she reviews the social scientific treatment of sexuality and the way it has shaped anthropology as a discipline. Weston argues that while anthropology often has taken the "flora and fauna" approach to sexuality, describing and categorizing "it" across various cultures, anthropology also has drafted sexuality into various debates that have defined it as a discipline, such as debates over nature and culture, the incest taboo, ritual and play, inheritance and work. Weston aims to treat contemporary sexualities in a similar way; she proposes to engage in the study of sexualities in order to explore larger theoretical questions.

For example, Weston takes up the gay migration from country to urban areas and winds up exploring "coming out narratives." She notices that in these stories the gay imagination is spatialized, creating sexual geographies that locate urban areas as the site of homogeneous communities of gay men and lesbians. When these areas are more closely observed, however, the presumption is undermined; differences among gay men and lesbians often trouble any sense of a homogeneous community. Still, the distinction of country and urban area persists in the organization of coming out stories, functioning as well to define homosexualities. For another example, Weston revisits the claim to family made by gay men, lesbians, and bisexuals in the 1980s and 1990s and finds herself drawn into the larger debate over whether this represents assimilation of an "alternative ideology" to a dominant discourse or whether it powerfully delegitimizes the naturalization of the family ties.

While her ethnographic approach to sexualities leads Weston to larger debates in social science, just as she had hoped it would, it also serves her resistance to queer theory. Weston asserts that the recent ethnographic treatment of sexualities has produced a rich variety of sexual practices (not all consistent with sexual identities); it, therefore, also has produced an anxiety over whether there is a sexuality prior to its various practices or whether sexuality is constituted as such in terms of these various practices. Weston concludes that recent ethnographic study of sexualities has produced its own version of queer theory, which is able to complicate and denaturalize the Anglo/Eurocentric conceptions of gender and sexuality.

While Weston seriously oversimplifies queer theory, ignoring its treatment of unconscious fantasy in relationship to dominant ideological discourses, her ethnographic approach does lead to other complexities, those produced when the ethnographer takes the role of native. Because anthropology's recent treatment of sexualities has depended on the ethnographic work of self-identified gay, lesbian, and bisexual anthropologists, Weston turns her concluding chapter into a reflection on the difficulties of doing ethnography of "one's own." Focusing on the compounded experience (or the hybridity) of being a native-to-be-studied and a professional-trained-to-study, Weston offers a chilling account of the rejection and marginalization facing anthropologists practicing various forms of autoethnography, even if they get jobs, tenure, and fame. But, she also argues that it is these anthropologists who will most likely retool social science for the 21st century. It is these anthropologists who, in embodying the contentious debates over scientific authority, are best able to elaborate strategies for self-reflexivity necessary to a critical social theory.

While *Fighting Words* and *Long Slow Burn* do not break new ground—at least not for those who have been close to the cultural criticism of the last 20 years, nevertheless, both are careful treatments of cultural criticism and its impact on social science. While neither is concerned to further elaborate the more radical aspects of the cultural criticism of the past 20 years, both focus on what matters most: the politics of race, class, gender, nation, ethnicity, and sexuality.

*Whitewash: Racialized Politics and the Media.* By John Gabriel. New York: Routledge. Pp. vii+219. $75.00 (cloth); $22.99 (paper).

Stephen Ellingson
*Park Ridge Center*

*Whitewash* is a provocative but maddening study about the social construction of "whiteness" in Great Britain and the United States. Gabriel examines the multifaceted and complex nature of white identity and outlines an ambitious research agenda to uncover how the mass media shapes racial politics and at the same time is itself used to further the ends of specific political groups. However, the book disappoints on sev-

eral levels: the case studies are underdeveloped, the methodology raises questions about the representativeness of the data, and conclusions about the effects of racist discourses are not grounded in reception data.

Fundamentally, the book is a study of how different versions of whiteness are constructed by the media and politicized groups in order to maintain, reclaim, or extend the power and privilege of whites. Underpinning his study is a crisis theory of change. He argues that a series of social and economic changes in the late 20th century, such as massive immigration, multiculturalism, the decline of nationalism in the West, and concomitant rise of transnational corporations, is undermining or at least threatening to undermine whites' control of the social, political, and cultural institutions through which they rule. This has mobilized the familiar racial "backlash" culture of the 1990s.

Gabriel explores the process by which whiteness is constructed and contested in several case studies of current hot-button issues: immigration control, policing, multiculturalism and political correctness, and white supremacy/separatist politics. The author is at his best when he carefully disaggregates whiteness into its varied forms, identifies the strategies different politicized groups use to create racial images and ideologies, and demonstrates how these different forms of whiteness are related. For example, in chapter 6, the author discusses efforts by far white groups to "[make] the defence of 'whiteness' an integral *and explicit* part of their political agendas" (p. 155; emphasis in original). After outlining the characteristics and goals of "white pride politics," which are self-consciously racist, he discusses how the discourses of British neo-fascists and David Duke's National Association for the Advancement of White People have been adopted and tamed by mainstream conservative politicians to propel their own, more subtle policies intended to restrict immigration or defang affirmative action. He argues that this co-optation of extreme right discourses is an attempt to reassert the hegemonic status of white identity and power (what Gabriel calls "normative 'whiteness'"), and thus naturalize or universalize whiteness. The author identifies the continuities between extremist and mainstream white discourses throughout the chapters. Each chapter contains several case studies, most of which are not developed in detail. They tend to serve as illustrations of his theoretical claims about whiteness rather than fully demonstrating how whiteness is constructed and why it is constructed in a particular manner.

Gabriel concentrates on the role of the mass media in the construction of whiteness, especially the tabloid press in Britain and talk radio in the United States. The media is both a forum to be manipulated by politicized groups to advance their racist agenda and an independent creator of new white discourses. Unfortunately, he fails to specify the conditions under which the media acts as an autonomous force in racial politics and when it is simply a tool to be baldly used by competing groups. Thus the media serves more as a source for data than as an explanatory variable. More troubling, however, is the nature of this data, which Gabriel never describes in detail. He does not discuss sampling procedures, selec-

tion criteria, or the process of data collection. Thus the reader has no confidence that the data is representative. Moreover, Gabriel does not identify his methods of analysis or provide the reader with reasons why she should believe his interpretation of the newspaper articles and excerpts from talk radio transcripts.

Another methodological weakness of the study is its inattention to reception. Although Gabriel explicitly focuses on the construction of whiteness discourses, he suggests that these discourses have effects, that is, they reinforce racial ideologies and maintain the power of whites. Yet without identifying how different audiences receive and act on these discourses, Gabriel cannot claim with confidence that various whiteness discourses had the alleged effects. Did anti-immigration discourses in the United States and Great Britain change how Americans and Britons think about race? If attitudes or policies changed, how can we know that it was the discourses and not some other cause or causes that brought about the changes? Gabriel does not answer these questions. In short, Gabriel's unsystematic use of media data and failure to make his interpretive methods and criteria public should make readers wary of his claims, despite their plausibility.

Theoretically, the book would pack a bigger punch if the author had discussed the broader implications of his findings for understanding how racism operates. Are the strategies by which whiteness is constructed unique to whites, or are they employed by dominant groups in nonwhite societies? Would this analysis look different if he were studying countries in sub-Saharan Africa, Southeast Asia, or Latin America? To be fair, this is not Gabriel's agenda, but it would enrich the study and would make it more than just another book that reveals the continuing presence of racism in the West.

*How Jews Became White Folks and What That Says about Race in America.* By Karen Brodkin. New Brunswick, N.J.: Rutgers University Press, 1998. Pp. xi+243. $48.00 (cloth); $18.00 (paper).

Ruth Frankenberg
*University of California, Davis*

This book's title must be taken literally; the reader should be forewarned that "becoming white" is a complicated process that has both pros and cons. Brodkin examines serial remakings of Jewish ethnoracial status and identity through the last 100 years of Jewish U.S. history. Through this period, Jews have, she argues, moved from the oppressive and subordinated status of "not-quite-white" to a more complex, ambivalently embraced status that might be described as for the most part quite white, but perhaps no longer quite Jewish—or at least not Jewish in the old way.

Brodkin reads the example of Jewish America against the broader framework of racial formation in which Jews, like all other Americans,

are emplaced. Here we come to the book's subtitle, *What That Says about Race in America.* In this text, Brodkin offers not only a rich revisioning of the location of Jews in the U.S. racial order but also makes forceful, provocative contributions to our understandings of the processes and ambivalence associated with Euro-ethnic assimilation into the racial category of whiteness. Further, in placing Jews at the center of her discussion, Brodkin is able to make subtler the sometimes rigidly bipolar mappings of racial formation that mark discussions of race in U.S. society.

Brodkin is interested in examining *both* how ethnoracial positioning has been imposed upon U.S. Jewry *and* how Jewish communities have conceived and shaped their own senses of self and identity in racial, ethnic, cultural, and political terms. Brodkin's description of her methodology for this study names it as "a slightly unorthodox combination of participant observation, insider ethnography and grounded theory" (p. 4). However, the text is also anchored in an examination of the ways racial formation, gender construction, and capitalist economics have been mutually constitutive throughout U.S. history. Brodkin locates Jews within a history of immigration and eventual assimilation and transformation, a process involving simultaneous shifts in that community's economic, racial, and sociocultural positioning in the United States.

Brodkin's "unorthodox combination" works extraordinarily well in this book. She contextualizes her insider's observations, insights, and memories (as well as those of her brother, parents, and grandparents) in a rich discussion of historical and sociological materials. Moreover, her combined strengths as an insider, ethnographer, and qualitative social scientist mean that she is able to come to new, often startlingly radical conclusions about the period and subject under consideration. For example, Brodkin notes that "part of my ethnic heritage was the belief that Jews were smart and that our success was due to our own efforts and abilities, reinforced by a culture that valued sticking together, hard work, education and deferred gratification" (p. 26). Brodkin contrasts this "belief in a Jewish version of Horatio Alger" (p. 26) with an analysis of the racially discriminatory workings of the GI Bill and the Federal Housing Authority in the years following World War II. In Brodkin's words, this amounted to "the biggest and best affirmative action program in the history of our nation, and it was for Euromales" (p. 27). Brodkin uses both childhood memories and sociocultural history to document this process and its impact on the United States social and racial landscape.

A crucial contribution of this text is Brodkin's argument that gender was central to the formation and reformation of the Jewish community, and to the marking of ethnoracial boundaries around dominant constructions of whiteness. Thus, for example, Brodkin writes cogently of how women in early 20th-century New York City were a crucial economic force both in the household and workplace, fought political battles over rent and wages, and, once married, enacted a vibrant sensuality and sexuality. All of this placed them clearly at variance with the dominant white construction of femininity.

However, Brodkin argues, after World War II, American Jewry began to benefit economically from a degree of inclusion in the dominant white world. At that point, an earlier hegemonic, socialist construction of Jewishness was replaced by a new hegemonic version: "a male centered version of Jewishness that was prefiguratively white, and a specifically Jewish form of whiteness" (p. 139). The earlier history of Jewish women's activism had thus to be rewritten and falsified. Ergo the trope of the "Jewish Mother" as a nagging and failed version of her white counterpart, an image that hints in a distorted way at the history of women in the Jewish community.

This text explores a broad range of economic, sociocultural, political, and ethical questions. In its conclusion, Brodkin argues that revisiting the earlier, Jewish socialist hegemony might provide tools with which to create more egalitarian, less oppressive cultural and political forms of Jewishness.

*Reading Public Opinion: How Political Actors View the Democratic Process.* By Susan Herbst. Chicago: University of Chicago Press, 1998. Pp. x+256. $16.00.

William Hoynes
*Vassar College*

In *Reading Public Opinion,* political scientist Susan Herbst provides an interdisciplinary look at the various and contested meanings of public opinion in contemporary American political life. As she sifts through her rich qualitative data, Herbst's underlying concern is with the relationship between conceptions of the often-slippery "public opinion"—how to know or measure it, when it matters, why it changes—and the "models of democracy" that correspond to such definitions of the public will. The result is an exploratory study of big questions, and Herbst manages to be both appropriately modest about her findings and theoretically sophisticated about the questions she asks.

The empirical core of *Reading Public Opinion* is a series of in-depth interviews with legislative staffers and journalists in Springfield, Illinois, the state capital, along with both a survey of and a handful of in-depth interviews with delegates to the 1996 Republican and Democratic national conventions. In all three cases, Herbst explores the way these knowledgeable and engaged political actors define and respond to public opinion. Given the complexity of the issues at stake, this is no small task. Herbst offers her respondents a series of open-ended questions, probes specific cases they raise, and asks them to reflect upon the relationship between mass media and public opinion. As a result, much of the book is Herbst's description of her data, chock-full of excerpts from the interviews, and her interpretation of what she terms "lay theories" of public opinion.

Herbst makes a persuasive case that these lay theories, in contrast to the scientific theories developed by opinion researchers, can have a significant impact on the workings of the democratic process. Indeed, since measures of public opinion are often wielded as a valuable political currency, political actors' assumptions about the sources and meaning of public opinion can shape both the policy-making process and the legitimate avenues for political participation.

Herbst finds that the three groups—legislative staff, journalists, and party activists—have quite different conceptions of public opinion. Perhaps most interesting is Herbst's finding that neither legislative staffers nor reporters rely upon public opinion polls. At least at the state level, on which this study focuses, issue polls are too expensive and infrequent, and they do not give much information about the intensity of public attitudes, so these political professionals find them of little, ongoing use. Instead of polls, legislative staffers turn to interest groups and media content to gauge public opinion, reasoning that these provide crystallizations of what informed citizens think. And journalists' definition of public opinion is based upon their "news sense" of the public mood, what Herbst describes as their own "imagined publics," and conversations they have with colleagues, friends, and news sources.

As she explores these lay theories, Herbst provides a nuanced reading of where these definitions come from and how they are connected to the daily constraints of her informants' occupations. And she highlights how such approaches differ from most academic research on public opinion, noting that "staffers and journalists find that their own indicators, which are clearly biased toward the more vocal and articulate constituencies among the public, serve them perfectly well" (p. 122). Citizens do exist on the radar screen of these political actors, but it is either organized citizens (to whom legislative staffers pay particular attention) or middle-class professionals (with whom journalists are likely to talk) that become the politically relevant public. Herbst insists that her informants do think about the public, but their definitions of that public produce an elite-oriented definition of the political process.

Herbst's party activists provide a different view of the public, one that is more consistent with a classical view of democracy. These activists are less inclined to value crystallizations of the public mood offered by interest groups or news reports, since they believe that public opinion is the aggregate of individual opinions, including those who are either unorganized or uninformed. In that spirit, party activists find polls to be useful measures of public opinion. More important, they believe that dialogue is the core of democracy, so the best way to know public opinion is through political talk with one's fellow citizens, or "local talk within social networks" (p. 179). Herbst argues that scholars know little about the ways people talk about politics in these social networks but that there is much to learn from more focused study of this microworld of political conversation.

Ultimately, this book raises more questions than it can answer, particu-

larly regarding the applicability of Herbst's state-level findings to the often poll-obsessed national political scene. But in raising complex questions about working definitions of public opinion and their relationship to a mass-mediated political world, Herbst has helped to map a research agenda and paved the way for further inquiry.

*Avoiding Politics: How Americans Produce Apathy in Everyday Life.* By Nina Eliasoph. New York: Cambridge University Press, 1998. Pp. x+330. $64.95 (cloth); $22.95 (paper).

Joshua Gamson
*Yale University*

I read this smart, energetic, slightly odd book about political apathy at the tail end of the Bill and Monica saga, when even the observation that citizens were uninterested in the impeachment trial elicited widespread disinterest. Read against Nina Eliasoph's concern for the sparseness and thinness of what she calls "public-spirited political conversation" (p. 14), what was striking about the recent scandal was how much people were talking about politics for once, even if it was just to say they wished people would shut up about them already. All the bored chat about public life underlines the important phenomenon Eliasoph tackles: how and why people "keep politics at arm's length in so many situations" (p. 10) and how and why, as she puts in it her central metaphor, American political discussion tends to evaporate so quickly and easily.

Many people have pointed out the impoverishment of American political discourse, but what distinguishes *Avoiding Politics* is not just that Eliasoph refuses simple explanations (television has made people dumb, for instance, or major institutions inhibit political engagement), or her facility with public-sphere theories, but that she actually went out and talked and listened. The book is the result of two-and-a-half years of ethnographic research with a range of civic groups in a "post-suburban" region: voluntary associations such as antidrug groups, a recycling center, and a high school parents' group; recreational groups such as a country-western dance club and a fraternal organization she calls "Buffalo Club"; and activist groups such as an antitoxics group and a peace vigil. While it exhibits some of the minuses of a converted dissertation—a tendency to pack in too much and to include tangential, so-and-so-might-argue discussions of too many other scholars—the book makes an innovative contribution to the important, ongoing discussion of American public discourse.

What Eliasoph found, most significantly, going to meetings, dance sessions, lodges, and in both her interviews and day-to-day conversations, was the puzzling, ironic disappearance of publicly minded talk, especially in more public contexts. It is not that people had no communal concerns but that they tended to speak about politics "backstage, in hushed tones"

(p. 16) and that "citizens' circles of concern shrank as they spoke in public contexts" (p. 6); at "each step in the broadening of the audience, the ideas shrank," she observed, such that "what was announced aloud was less open to debate, less aimed at expressing connection to the wider world, less public-spirited, more insistently selfish, than what was whispered" (p. 7). The core argument of the book is that the kinds of political disconnections and silences so often bemoaned by critics are actually rooted in the contexts available for political *conversation*. The "informal etiquette" in her field settings, for one thing, "made some political intuitions speakable, and others beyond the pale of reasonable, polite discussion" (p. 7). Indeed, ideas about what counts as "political" change from one context to the next, and in the American public sphere, this leads to a paradox: what marks a context as "public" is "often precisely the fact that the talk there is so narrow, not at all public-minded" (p. 230).

The rest of the book is a heavily detailed, closely observed account of the way talk does and does not work in Eliasoph's various settings, along with a chapter on the role of local newspapers in the "cycle of political evaporation"—which picks up, not incidentally, on an explicit, critical, and insufficiently considered link between local conversational contexts and the larger structures in which they are embedded, be they media structures, gender structures, or governmental ones. The level of detail, which is the book's strength, is also at times its deficit: unsurprisingly for a book that seeks to focus attention on talk, *Avoiding Politics* is quite talky, and not always in ways that further clarify; while it is wonderful to hear the many voices of her subjects joining the author's, sometimes the conversations recounted simply contribute banality rather than revelation.

Nonetheless, using solid qualitative research coupled with sharp, quirky insights, the book poses terrific challenges to those interested in the workings of political discourse, political belief, and the public sphere. *Avoiding Politics* makes a strong case for the everyday effort that goes into constituting and marking off a "public sphere," the ways that enterprise varies from place to place, each with its own set of talking norms, so many of which uninvite, or slowly vaporize, politically minded conversation. "Apathy takes work to produce" (p. 6), Eliasoph argues, and her book is a big, innovative help in the ongoing attempt to think and rethink strategies for producing something else, for creating spaces in which "public" talk is neither trivial nor inhibited.

*Governing Out of Order: Space, Law and the Politics of Belonging.* By Davina Cooper. London: Rivers Oram Press, 1998. Pp. xiv+242. $50.00 (cloth); $19.50 (paper).

Neil Brenner
*New York University*

Davina Cooper's most recent book is composed of a series of essays framed by a broader methodological argument regarding "political authority and its boundaries at the end of the twentieth century" (p. 3). Drawing upon Foucauldian political analysis, discourse theory, and cultural geography, Cooper attempts to reconceptualize the traditional state/civil society and public/private distinctions in terms of "bodies that govern" (p. 8) and "governing techniques" (p. 11). The book elaborates this theoretical agenda through case studies of diverse struggles over the boundaries of national and local governance in post-1970s Britain. The book is directed primarily at a poststructuralist readership but will also be of interest to scholars in the fields of state theory, organizational sociology, legal studies, and local government studies.

Cooper focuses on the issue of "governmental excess"—how institutions exceed their legally delineated roles, how individuals and organizations construct the boundaries of political authority, and how those boundaries are in turn contested through social conflict. Even in an era of neoliberal deregulation, Cooper argues, the question of governmental excess remains urgent insofar as the boundaries of state authority are contested in diverse realms of social life. Drawing upon recent Foucauldian scholarship, Cooper distinguishes three modes of governing—"direct governing" (based upon a visible form of rule), "governing at a distance" (based upon internalized norms), and "mid-way governing" (based upon a combination of the latter; pp. 12–13). Cooper argues that each of these forms of governance plays an important role in conflicts over political authority. Cooper endorses an antiessentialist conceptualization of state power as a "contingently articulated set of identities, with no fixed form, essence or core" (p. 9). In addition to these claims, Cooper argues for an interpretive framework that is attuned to the spatiality of governance practices. Finally, Cooper maintains that the issue of "belonging"—membership within a political community—is also a major stake in conflicts over governance.

Cooper explores these themes through case studies of various "flashpoint moments" (p. 167) of conflict over governance in contemporary Britain. Chapter 2 elaborates a discourse analysis of British parliamentary debates during the Canada-Spain fishing rights dispute of 1995, in which diverse conceptions of British national identity were articulated. Chapter 3 investigates debates on religious education in British schools provoked by the Christian Right during the late 1980s. Chapter 4 analyzes the discursive construction of fiduciary principles, taxpayers' duties, and local community during the 1980s. Chapter 5 examines the conflicts

that erupted between local state institutions and neighborhood activists when the head teacher of a London school criticized a ballet version of Romeo and Juliet for its heterosexist bias. Chapter 6 examines the struggles of Jewish residents to construct an "eruv" (a symbolic perimeter used during the Sabbath) in a London neighborhood. Chapter 7 discusses struggles over hunting rights in the British countryside. In the final chapter, Cooper defends a "transgressive" (p. 181) normative position as a means to politicize, and thereby to unsettle, hegemonic discourses and institutional arrangements.

In each of the aforementioned chapters, Cooper attempts to relate the specific topic under investigation to the broader themes of the volume: "The common ground may appear obscure," Cooper writes (p. 4), but "together these issues offer a mosaic of modern, liberal governance." However, after a careful reading of the case studies, their "common ground" remained opaque to this reader. Despite the lucid presentation of each chapter, the design of the broader "mosaic" in which they are situated is difficult to discern. Many of the chapters were originally published as articles, and their inclusion within a single volume leads to an eclectic juxtaposition of diverse "issues" and "episodes" (p. 21), which are said to instantiate the book's general themes. In short, this is a book whose parts are greater than the whole into which they have been subsumed.

Because Cooper focuses above all on the episodic time of small-scale events and struggles, the macroinstitutional context in which they are situated—namely, Thatcherite Britain—is largely bracketed. Indeed, despite her antiessentialist epistemological stance, Cooper's analysis of discourses of governance often veers toward an idealistic representation of political conflict as a clash of cultural meaning systems. The institutional materiality of governance is thereby rendered nearly invisible. Finally, Cooper tends to reduce the spatiality of governance to its locational specificity—the fact that conflict is always situated within particular locations. Consequently, the historical and geographical specificity of contemporary transformations of governance, as analyzed by many of the critical geographers cited in the book, is not adequately theorized.

Despite these deficiencies, however, Cooper's book constitutes a well-written and creative contribution to contemporary debates on changing forms of governance in western Europe. While the individual case studies are not coherently integrated into a synthetic whole, they are nevertheless illuminating and interesting in their own terms. More important, the opening chapter contains a number of methodological prescriptions that may yet be mobilized to push studies of governance in new and unforeseen directions.

*Democracy in Translation: Understanding Politics in an Unfamiliar Culture.* By Frederic C. Schaffer. Ithaca, N.Y.: Cornell University Press, 1998. Pp. xvii+168. $39.95.

Jeffrey W. Rubin
*Rutgers University*

In this elegant and lucid study, Frederic C. Schaffer asks what democracy means to people in Senegal. By combining detailed interview material with analysis of language and culture, he shows first that the French-speaking elite disagrees among itself about the meaning of *démocratie*. For those who support the party in power, *démocratie* is an unfolding process that should be judged by the presence of electoral competition and the stability and vitality of the country. For those in opposition, *démocratie* tends to be equated with genuinely fair elections, as well as the acid test of alternation of the party in power, which has not occurred in Senegal.

Schaffer then examines the ways this elite presents its understanding of *démocratie* to the majority Wolof-speaking population, how the meaning changes "in translation" for political and cultural reasons, and how the responses of Wolofones in turn modify the uses of *démocratie* among the French-speaking elite. Thus, when opposition leaders speak to Wolofones, they emphasize the authority of the people rather than institutional arrangements. Schaffer shows how opposition politicians draw on the negative associations attached to the Wolof words for sovereign (*burr* or *nguur*), as well as how the words for selection and deposition of rulers, *fal* and *folli*, correspond more to deliberation and consensus than to voting. "These projections and metaphors," Schaffer concludes, "carry with them meanings embedded in popular culture that pull *démocratie* from its semantic foundation" (p. 52).

Schaffer then turns to the multiple meanings of *demokaraasi* among the Wolof-speakers. He shows that *demokaraasi* means, at various times and places, consensus, competition, community solidarity, and a share in the wealth and pleasures of the elite. In contrast to Western democracy's emphasis on autonomy and participation, *demokaraasi* is more concerned with welfare and evenhandedness. Thus, when Wolofones vote, they may be acting to affect public policies but also to gain resources for survival or maintain the cohesion of a neighborhood, village, kinship, or religious group.

In analyzing the meanings of *démocratie* and *demokaraasi*, Schaffer looks closely at institutions, cultural practices, and language in historical context. This is the strength of the work, and it marks out a new path in political science—one that finds common and fruitful ground with cultural studies. There is serious attention here to language and culture, not just as discrete variables, but as the complex, constructed, and changing phenomena out of which politics is woven. For example, in analyzing why some voters comply with religious directives, or *ndigal*, with regard

to how they vote, while others do not, Schaffer concludes that "the inten-
tions of Senegalese voters are many-sided. In addition to calculations of
public good and private interest, these voters must weigh considerations
of religious devotion, community cohesion, and material exchange" (p.
114). One might even suggest—as Schaffer does not—that how an indi-
vidual acts in such a situation is less a "rational choice" than a complex
individual and cultural product

Schaffer frames his analysis of *démocratie* and *demokaraasi* between
two succinct discussions. In the first chapter of the book, he illuminates
in very short order the complex and potentially muddy topic of democ-
racy—its U.S. or Western meaning, that is—and how to evaluate it.
Schaffer argues that outside the United States and Europe, it is a mistake
to assume that electoral competition brings about accountability, which
he sees as the key substantive achievement of democracy. In the penulti-
mate chapter, after his incisive tour of Senegalese meanings, Schaffer re-
turns to this question of U.S. democracy. Having identified the absence
of accountability in Senegal's electoral system, Schaffer asks, how might
it become more democratic? In response, he touches briefly on such mat-
ters as civic education, economic security, and the secret ballot.

These two "democracy" chapters underscore both the strength of
Schaffer's prose and the limits of his framework. As someone who has
migrated from political science to anthropology and history, I would say
that he has drawn a too-tidy border around his rich material and analysis
in order to package it for political science. Much of what he says about
Western democracy and the Third World, though elegant, is obvious, and
his comments on democratization are entirely speculative. Schaffer suc-
ceeds in saying, to people who make blanket assertions about the demo-
cratic character and salutary benefits of elections, that democracy, when
translated, is not necessarily what they think it is. That is his aim and
achievement and arguably a quite valuable contribution.

What more might he have done? He might have used his nuanced ac-
count of meanings to illuminate the workings of power in Senegal. This
would have meant showing how daily life, politics, and power relations
affect each other over time. Schaffer might also have problematized the
notion of democracy in the United States more than he does, raising
deeper questions about competition and accountability at home. While
taking on U.S. democracy and Senegalese politics in this fashion would
have been a large task and a longer book, Schaffer's rich interview mate-
rial and keen analytic ability indicate that he is well suited to the task.
Had he undertaken this, the result would have been more an anthropolo-
gist's or historian's book than a political scientist's, but it would have
told us also more about politics.

*Making Race and Nation: A Comparison of South Africa, the United States, and Brazil.* By Anthony W. Marx. New York: Cambridge University Press, 1998. Pp. xviii+390. $29.95.

David Theo Goldberg
*Arizona State University*

Anthony Marx has written an analysis of the role played by racial configuration in nation making in the tradition of Pierre van den Berghe's and George Frederickson's comparative work. Marx is concerned to account for the apparent differences played by race in the national construction of the United States, South Africa, and Brazil. The central thesis is straightforward and simple. Political elites resorted to racist exclusion, most notably in the form of de jure segregation, to consolidate whites in the face of intrawhite conflict (Civil War in the United States, the Boer War in South Africa), national instability, and potential demise in power. So de jure segregation emerged to unite whites in these societies. By contrast, Brazil suffered no internalized conflict among whites, and so no need to resort to segregation of blacks as a way of uniting a divided nation identified with whiteness. In Brazil, discrimination accordingly assumed less overt forms. In this scheme of things racial formation is taken to be imposed more or less topdown by elites seeking to ensure solidity in their nation building in the late 19th and early 20th centuries. Marx undertakes to render this thesis more complex and nuanced by arguing that formal exclusion prompted resistant racial identities among people of color that were necessary in turn for protesting such exclusion and mobilizing for inclusion and resource sharing. Here again Marx differentiates Brazil from the other two instances, for in the former, lack of formalized racism is deemed to result in the relative lack of resistant race-based identity formation.

Marx silently takes for granted a number of problematic theoretical assumptions underlying his account. Thus he undertheorizes the state, race and racial formation, and the nature of politics, not least in light of considerable recent work concerning all three. First, there is almost total absence here of any account of the nature of the state. Marx conceives the state minimally and traditionally in a Hobbesian vein, as acting instrumentally and coherently to ensure and maintain social stability and security, reduced to the unified actions and plans of elites. The state is seen simply to act and to use race instrumentally to the ends of stability and security (pp. 4, 13).

It follows, second, that there is no account of race and race making in the book beyond what elites and resisters are taken superficially and obviously to do in relation and response to each other. So Marx offers no account of how race is used, what it stands for materially and symbolically, what work and conditions in different contexts it is able to effect beyond the bald unification of whites in the face of their own potential conflict. Accordingly, he suggests a totally reductionistic sense of racial

formation and discourse, determined by a mix of economics and politics, the effect of which is to force an artificial similitude between the United States and South Africa in order to save the thesis. Superficially both the U.S. Civil War and the Boer War were conflicts for control over territories and wealth. By contrast, however, the Civil War was not an *ethnic* conflict among whites that necessitated state imposition of segregation to resolve. Indeed, in at least the South African and U.S. cases the state was involved in racial formation and exclusion all the way down, in and from the very constitution of the respective social formations. Thus the deeper question concerning racial formation is not simply what accounts for the obvious differences between the three societies under discussion but why whites historically and contextually see themselves as needing to be united on specifically *racial* grounds. And the posing of this question suggests that it is not just a function of the fear of political and economic losses but also a mix of fear and exploitability of black worlds. Thus Marx fails to consider that the modern state, both in its theoretical conception in the likes of Hobbes, Locke, Rousseau, and Hegel, and in its material instantiation, is predicated on racial design and racist exclusion. Race is constitutive of and not just instrumental to modern state formation. It is the suggestion of white anxiety about black control then that ties the emergence of segregation to the history of racializing discourses and state formation preceding the late 19th century. About this Marx says nothing.

In similar fashion, Marx has a very narrow sense of politics as mobilizing the electorate for the sake of gaining material resources. This flattens the sphere of the political, tearing apart the politics of culture from the culture of politics and from the role of culture in political transformation (cf. p. 261). It also reduces politics to material mobilization rather than for seeking power and power sharing for their own sakes, thereby failing to raise the question of the articulation of power in and through racial configuration. In short, race is not peripheral to state formation in modernity, an afterthought to effect state security, but constitutive of modern state power, a principal arrangement in terms of which modern social power is articulated.

In a book replete with questionable fact and more questionable interpretation yet, most damaging is the assumption that black protest and "direct action," in both the United States and South Africa, only followed and was a consequence of white action. Thus Marx claims that black Americans began to engage in direct action only after *Brown v. Board of Education* in 1954 (p. 148) and that black South African protest emerged with volatility only in 1960 (p. 107). Marx reduces the forming of a black identity in both societies to a reaction to white state definition and control, failing to understand the investment in black culture for its own sake (cf. pp. 264, 266; consider the value of African retentions in black culture in the United States and Brazil). Marx's assumptions belie the considerable resistant action led by blacks in the United States and South Africa long predating these events.

In *Making Race and Nation,* then, Anthony Marx has written a work of superficial synthesis, borrowing a thread here and a stitch there, purporting to cover considerable historical and analytic ground but only at the cost of any body of literature; the overarching thesis, purporting to account for the differences in historical experience concerning the relation of race to nation-making, is forced at best—misguided at worst. The singular virtue of the book, then, remains in the fact that in the face of considerable theoretical silence about race and the state Marx at least has raised the question.

*Institutional Design in New Democracies: Eastern Europe and Latin America.* Edited by Arend Lijphart and Carlos H. Waisman. Boulder, Colo.: Westview Press, 1996. Pp. xvi+265. $69.00 (cloth); $18.95 (paper).

Geoffrey Pridham
*University of Bristol*

Institution building forms a central part of regime change. In democratization, it is quite crucial to the transition stage with important consequences for eventual democratic consolidation. This is for a number of reasons. The constituent process signifies the kind of system its creators would like it to be, while indicating pertinent ways in which the new institutions differ from those of its discredited predecessor, especially by introducing constraints on executive power. Moreover, the way of deciding on the constitution—if not some key elements in it—attests to the ability of elites to foster consensus behind the new and still fragile settlement. The form of the system established will clearly affect its performance potential and, hence, also the prospects for its own legitimation.

Despite this, institution building has not received much attention in theoretical or comparative work on democratization. Often, it has featured indirectly within other concerns, such as elite pacts and political choice. Now, the recent system changes in Eastern Europe have highlighted this question, while—as we learn from this volume—it has acquired a new emphasis in work on Latin America. In addition, the former region—as well as some Latin American examples—have drawn our attention to the concurrence of political and economic transformation and also the relevance of institutional arrangements in the latter respect. This volume is therefore particularly welcome.

In specifying institutional design, the editors focus on electoral systems, executive-legislative relations, and mechanisms of the market economy. These are seen as substantive issues relating to system change. All the chapters are indeed competent and informative: some are better are than others, while there are certain problems with the overall comparative approach of the book. The best single comparative contribution is that by Barbara Geddes, which looks at the choice of political institutions in both regions in a systematic way. She relates her theme to broad contextual

factors such as party systems and cultural legacies as well as economic transformation. Eva Voszka's analysis of privatization is useful and clear, although confined to three countries in Central Europe. The concluding chapter by the editors on the results of their collective book is lucid and insightful. Unfortunately, it is made necessary because the broad themes are not as a whole explored enough in the individual chapters.

Interregional comparisons of this kind are now well-established in the literature on democratization, whether these relate separate parts of Europe (in different periods of regime change) or one of these (more often Southern Europe) with Latin America. But there is a narrowness in the empirical base of this comparative volume. Apart from four comparative chapters (two cross-regional and two intraregional), there are eight case studies. Three are intended to cover Eastern Europe (two on Poland and one on Hungary), and there are five on Latin America (two on Argentina, one on Brazil, one on Chile, and one on Mexico). Thus, there is some imbalance between the regions, while the treatment of Eastern Europe is confined to part of Central Europe, and the Balkan countries are ignored. One feels that the subtitle should read: "Central Europe and South America." This problem of scope also casts some doubts on the comparative results of this volume.

By and large, the case studies are, whatever their individual merits, not written in comparative perspective, while several are rather too descriptive. There is no real attempt to explore the interaction between political and economic transformation throughout the book, although some of the contributors make passing mention of this. This overarching theme, focusing on institutional choice, could have been exploited by the editors by encouraging their able team to develop their chapters in this direction. That would have much enhanced the merit of this work. However, a number of other familiar themes receive attention in a variety of chapters, including the relationship between party and electoral systems and the preference for a presidential or parliamentary system. On the former, nothing new is stated, for the relationship is complex (there are various factors involved), so PR cannot be held responsible for party system fragmentation. On the latter, there are some interesting empirical insights. In effect, a tendency for semipresidential structures is noted. Other important structural features of new democracies, such as center-periphery relations, are ignored. On the other hand, many of the authors emphasize the importance of political and institutional legacies, a theme that deserves more general attention in democratization studies.

*Symbol and Ritual in the New Spain: The Transition to Democracy after Franco.* By Laura Desfor Edles. New York: Cambridge University Press, 1998. Pp. xiii+197. $59.95 (cloth); $19.95 (paper).

Jan Kubik
*Rutgers University*

During the last several years, the third wave of democracy—as Huntington christened it—became a subject of countless analyses. Not much, however, has been written about the cultural dimension of democratization. Moreover, in this meager literature, one strand has become dominant. Its central category is political culture; its methodology—public opinion surveys. The works that deal with the politics and substance of cultural transformations accompanying democratization, not just people's opinions, or with the formation of new political symbolism by the ascending elites, not with people's understanding of this symbolism, are few and far between. That is why Laura Edles's book is so important, and that is why it deserves both wide readership and close scrutiny.

The book delineates and analyzes the cultural field within which the Spanish democratization occurred. Analytically, Edles relies mostly on cultural sociology, and semiotics (Umberto Eco) and cultural anthropology (Victor Turner). Her methodology is discourse analysis of crucial text produced during the Spanish transition; her technique is content analysis of key newspapers and magazines.

The book has three parts and nine chapters. Part 1 contains an overview of various theories of democratic transitions, an exposition of Edles's own theory, and a brief introduction to the political history of Spain. Part 2 provides a detailed analysis of the symbolic dimension of the Spanish transitory politics. Part 3 contains an analysis of the 1978 Constitution-making process, a brief discussion of the Basque exceptionalism, and conclusions.

The main interpretive strategy is to identify the core symbols (or representations) used by the Spanish elites and to demonstrate how these particular symbols came to dominate the political imagination of the most Spaniards. The four core symbols, identified by Edles, are "a new beginning," "Civil War," "national reconciliation/convivencia," and "democracy." Edles argues that without the hegemonic culture of reconciliation, organized around these symbols, the Spanish peaceful and successful democratization would not have been possible. The second strategy, borrowed from anthropology (most prominently from Victor Turner), is the conceptualization of the transition as a rite of passage, composed of the three stages: separation, liminality, and reaggregation. The 1977 elections are the core event of separation, the Moncloa pacts stand in the center of the liminal phase, and the drafting of the 1978 Constitution opens the way for reaggregation.

At a high level of abstraction, the author's analysis is convincing: the Spanish transition is usefully construed as an accumulative process, with

each major event seen as a tipping point in the transformation not just of the political field, but also the society's (political) culture. What I missed, though, was a more detailed description of the actual political-symbolic battles politicians undoubtedly fought to achieve hegemony over the cultural domain.

This omission is not accidental; it is typical of the structural analytical frame the book employs and may be seen as its major shortcoming. The principal tool of structural analysis, the dissection of narratives, representations, and so on into sets of binary opposition, is used with abandon. It helps to understand the basic contours of the Spanish discourse on transition, but one is left wondering whether all arguments and positions were usefully reducible to the elegant logic of binary oppositions. Structural analyses (of texts) are particularly convincing when they are complemented by the processual analyses of (inter) actions (political praxis). It is a pity that the latter type of analysis is far less developed than the former.

The principal binary opposition deployed in the book is that between the "sacred" and "profane." The symbols and realities of the "new beginning," "democracy," and "national reconciliation" belong to the sacred domain; "civil war" is the principal symbol of the profane. At times, this conceptualization is illuminating, but its relentless application to almost all events, processes, actors, and so on is worrisome. One begins to suspect that, as with any "total" structural binarization, also here the comprehensive neatness of omnipresent dualisms is more an artifact of the author's skill than an inherent feature of the analyzed reality. There are also problems with the way the opposition between the sacred and profane is set. In most cases, this opposition means "good/bad" (example: "national reconciliation/Civil War"). But Edles uses sometimes the more conventional meaning of the "sacred/profane" opposition, that is, "extraordinary or holy/mundane" (most importantly on p. 93). This meaning is more useful, for it underscores the fact that such complex political/cultural processes as the Spanish democratization are multileveled. Whereas the symbolic wars or compromises between the forces of "good" and "evil" (relative, of course, to the actors' own viewpoints) transpire at the level of the "sacred," the whole gamut of deals or confrontations occurs simultaneously at the more "strategic" (Edles's term, p. 93) levels of the "profane."

The field of studies on the cultural/symbolic dimensions of politics is relatively barren. Thus, despite some criticism, I recommend Edles's book wholeheartedly. After reading it, I have not only a fuller grasp of the Spanish transition to democracy, but also a better understanding of the tremendous role cultural factors play in such processes.

*The New Language of Qualitative Method.* By Jaber F. Gubrium and James A. Holstein. New York: Oxford University Press, 1997. Pp. x+244. $49.95 (cloth); $24.00 (paper).

David Silverman
*University of London*

Textbooks on qualitative method are frequently strewn with false oppositions. Either they are dull nuts and bolts affairs or they purvey a particular theoretical paradigm and dismiss (or exclude) everything else. By contrast, this book integrates theory and method and sites itself in the crucial debate about reality and representation in qualitative research.

The first half of *The New Language of Qualitative Method* reads as a very helpful textbook introduction to four different idioms of qualitative inquiry, each of which is illustrated with detailed examples. Naturalism is derived from the Chicago School's injunction to "get out there." It is reluctant to impose meaning and seeks to assemble casts of characters with their own stories.

Ethnomethodology takes naturalism's attention to the detail of everyday life but locates that life in talk-in-interaction. Emotionalism (a new and provocative term) desires more intimate contact with research subjects and favors the personal biography as both topic and resource in its project of displaying researchers' "feelings about feelings." Finally, postmodernism takes emotionalism's interest in narrative experiment but weds it to the deconstruction of the tropes of the subject and the field. So far so bad. For, as Gubrium and Holstein forcefully argue, if we stay entirely within the boundaries of each idiom, qualitative researchers become organized into rival armed camps, unable to learn from each other or to think laterally beyond a theoretical impasse.

At this point, the authors move beyond a helpful but descriptive student textbook to develop a passionate argument from which we can all learn. For Gubrium and Holstein, qualitative researchers inhabit the "lived border between reality and representation" (p. 102). On this border, in their view, each idiom veers too far to one side. So naturalism's pursuit of the content of everyday lives offers us deep insights into the "What?" of reality at the price of the "How?" of representation. By contrast, ethnomethodology, they claim, gives rewarding answers to "how" questions but underplays the "what" of contextual givens. Emotionalism claims to deal with both reality and representation but only by reifying emotion. Finally, postmodernism's focus on representation, they suggest, can lead to a nihilistic denial of content.

However, this book is ultimately not about that kind of "yes, but" style of intellectual history. Instead, the authors offer three valuable practical ploys for the qualitative researcher.

First, seeking a middle ground to "manage the tensions between reality and representation" (p. 114), Gubrium and Holstein show how we can give voice to each idiom's silenced other. The figure of the insider, so

dear to naturalism, can be treated as "a represented reality" (p. 103), as can emotionalism's feeling subject. Equally, conversation analysis's (CA's) account of institutionality and Sacks's membership categorization machinery show, for Gubrium and Holstein, how ethnomethodology can put meat on the bare bones of representation. Last, while we must respect what postmoderns tell us about the play of signifiers, this, they argue, can be treated as an incentive for empirically based description, not as its epitaph.

Second, the book shows the way in which how, what, and why questions need not be mutually exclusive if we are prepared to use analytic bracketing to focus on one question at a time. So, for instance, CA can give us sophisticated answers to the how question by focusing on the artfulness of practical reasoning. The contexts that such reasoning uses as resources can then be examined by institutional ethnographers' descriptions of local cultures fashioned for specific goals and audiences.

Finally, Gubrium and Holstein reawaken qualitative researchers to inventive ways of thinking about why questions, which derive more from the postmoderns than from structural sociology. For instance, social constructionism can be treated less as a theoretical position and more as a set of contemporary institutional practices in which "people-defining organizations" (p. 204) survey and deprivatize subjects and, in so doing, create multiple, hyphenated subject positions.

Of course, this is hardly noncontroversial. For instance, the authors' characterization of ethnomethodology's project will inevitably raise a few eyebrows among practitioners. Again, for my taste, they are far too kind to emotionalism, which, as I have argued elsewhere, has a lot to answer for in its siren calls for the authentic voice of the subject (Paul Atkinson and David Silverman, "Kundera's *Immortality:* The Interview Society and the Invention of the Self," *Qualitative Inquiry* 3 [1997]: 304–25).

However, these are mere quibbles that should not mask this book's real achievement. In an accessible way, Gubrium and Holstein have contributed to breaking the intellectual stranglehold of both the narrow-minded methodological know-nothings and the theoretical high-flyers who would not know a datum if it hit them over the head—no mean achievement.

*Changing Families: An Ethnographic Approach to Divorce and Separation.* By Bob Simpson. New York: Berg, 1998. Pp. xiii+181. $55.00 (cloth); $19.50 (paper).

Joseph Hopper
*University of Chicago*

When students, friends, and my parents ask what the difference is between sociology and anthropology, I rarely muster a good answer. The substantive, theoretical, and methodological boundaries between the two

seem less distinct than they might have once been. *Changing Families* reminds me, however, that sociologists still have much to learn from anthropologists—not because anthropologists are engaged in a distinctly different enterprise, but because they are steeped in a way of thinking and in a theoretical tradition that illuminates social phenomena in ways we too often miss.

The study, the data, and the writing in this book all look like standard qualitative sociology. Simpson spent nearly a decade in various urban settings in England and Wales, working with other social researchers (sociologists, lawyers, economists, and social policy analysts) on decidedly policy-oriented research. The research began in 1985 with an effort to evaluate divorce mediation, which involved surveying 788 individuals who were going through court or mediation procedures. Of these individuals, 151 were then interviewed using qualitative methods, and many were interviewed or surveyed again in various follow-up studies that continued until 1991. With the policy goals of the research completed years ago, Simpson returns to the qualitative data to ask bigger questions about family and divorce.

The questions he asks are about kinship—a topic that should be central to sociology but rarely is in practice. How is it that relations are defined, boundaries set, decisions made about who is "in" and who is "out"? How are these relations maintained, performed, and internalized? How, indeed, does all of this happen when the basic "generative unit" of kinship in our culture—namely marriage—gets reconfigured via divorce? Simpson's particular focus is on the bewildering complications that divorce presents for divorcing spouses who are parents. Parents and children in our culture are bound by "blood," and their kinship is irrevocable. What happens, then, when parents try to sever the bond of kinship with each other while preserving such kinship with the children they share?

In a fascinating chapter about postdivorce financial arrangements, for example, Simpson argues that difficulties in arranging for child support are partly a result of conflicting efforts with regard to kinship. In the tradition of Malinowski and Mauss, he reminds us that money and gift exchanges are rarely just practical matters; they are rather laden with meanings that help define relationships. Money and gifts within families *express* our kinship'—they show commitment, express feelings, and thereby reinforce familial interdependency and social solidarity.

What happens when parents divorce, then? First, parents try to cut ties with each other via a clean financial break that expresses appropriate distance and impersonality. But second, they wish to provide for and maintain close ties with their children. For a father facing the typical scenario where his children now live with their mother, these two efforts come into conflict. To provide for his children, he must funnel money through the mother since it is she who maintains the children's everyday well-being. This not only establishes an ongoing direct exchange with his ex-wife (which he is trying to sever), but it also robs him of the meaningful direct exchange with his children, an exchange that would help estab-

lish symbolically a relationship of "depth and duration" (p. 125). Many fathers therefore resist making payments to their ex-wives; they try to link payments with more direct access to the kids and more control over how it is spent, or they instead spend the money on special gifts that they give directly to the kids. Mothers, not surprisingly, feel frustration and resentment when fathers do not pay (which should consist of "regular, quantified, market-like contributions," p. 125), and, well attuned to the symbolic meaning of money, they interpret financial delinquency as fathers not caring about the children's well-being. In short, Simpson shows that behind the economics of divorce lies a meaning-rich world where the "vague calculus of family obligations . . . [is] made explicit in the process of dismantling and reconfiguring family and household" (p. 110).

This chapter, which was published previously as an article, is the highlight of the book. Other chapters consider conflicts about parenting itself, the difficulties in parenting alone in a culture where parenting roles are usually complementary, and the difficulties in disentangling from one's ex-spouse when kinship structures have been so strongly defined via the severed bond of marriage. Each of these chapters, too, stands alone. Not all were as appealing to me as the chapter on exchange and child support, but I appreciated most for initiating a way of thinking about divorce that has surely deepened my understanding of it.

The value of this book for sociologists, then, is not that it contains enumerable findings that we can add to our storehouse of knowledge, but that it provides a different way of thinking about what constitutes the social world of family. And it brings to the topic a rich set of insights and ideas about the very same phenomena we sociologists are working to understand.

*Growing Up in Stepfamilies.* By Gill Gorell Barnes, Paul Thompson, Gwyn Daniel, and Natasha Burchardt. New York: Clarendon Press, 1998. Pp. viii+336. $77.00.

Linda J. Waite
*University of Chicago*

*Growing Up in Stepfamilies* uses intensive interviews with 50 adults who entered stepfamilies between the ages of 7 and 16 to examine the impact of this experience on the childhood and adult lives of these individuals. The respondents had all been surveyed as part of the National Child Development Study, which attempted to follow all 17,000 children born in the United Kingdom in a single week in 1958 and used structured survey questionnaire interviews. Those interviewed for the study of stepfamilies were selected from among those who became stepchildren by age 16. The lengthy and detailed interviews with these stepchildren—now adults—form the basis for this book.

The book begins with a discussion of the theoretical framework that

undergirds the study—a combination of the life story and family systems approach. The life story approach explores "the evolution of the present and recent past through in-depth retrospective interviewing" (p. 28). These interviews are shaped by the interaction between the interviewer and respondent, allowing the researcher "to reconstruct typical patterns of behaviour, subjective interpretations of experience and accounts of feeling, and unconscious assumptions conveyed by the narrative form" (p. 28). The family systems approach attempts to arrive at a contextual understanding of the behavior on individuals; it examines psychological development through feedback processes in intimate relationships and across different levels—between individuals, families, and communities.

The authors organize the book along dimensions that the interviews revealed to be important. They begin by looking at communication and roles in both the original family and the stepfamily. Families that are broken apart and reformed go through a period of loss and realignment, according to the authors, and children must reach an understanding of these changes. Relatively few children in this study felt that their parents did an adequate job of explaining the situation to them. Some parents— in fact many parents—gave their children little warning about these major changes in their lives, sometimes telling them simply that their mother was gone or moving a new spouse in without saying anything at all about it. Some lied or deliberately misled their children about who their biological father was, where he lived, or other important information. This failure colored the child's accommodation to the new arrangements and feelings about the stepparent.

The loss or absence of a parent deprives children of a valuable resource and poses a threat to the acquisition of a secure identity for the developing child. The authors focus on the factors that allowed some children to deal successfully with this loss and others to suffer long-term harm. Death of a parent seemed to be handled especially poorly by both families and social institutions such as schools, making it difficult for children to grieve. Most children whose parents divorced lost contact with the absent father. Extended family members sometimes pitched in, as did neighbors in some stable neighborhoods.

Gender plays an important role in stepfamily relations; the stepchildren in this study seem to hold stepmothers to much higher standards than stepfathers so that more stepmothers are judged quite harshly. In Britain in the period studied here, stepfathers who provided for the family financially and stayed out of the way were generally seen as a success. Stepmothers needed to do much more to be successful and more often failed.

The book also focuses on the impact of having lived in a stepfamily on the lives of the children. The adult stepchildren seemed to have done well occupationally, with many having risen from their family of origin. But many felt the long-term consequences of their childhood experiences in their own relationships. The stepchildren whose parents had divorced were much more likely to have divorced themselves than stepchildren whose parent had died, for example.

Family disruption was quite uncommon in Britain in the 1960s and early 1970s, when most of these children entered stepfamilies. The social stigma attached to family disruption and to stepfamilies shaped the behavior of the parents and children in this study. Divorce and unmarried childbearing have both become much more common since that time, so that the experience of the stepchildren who appear in this volume was likely quite different from the experience of children who are members of stepfamilies today.

The perspective of development psychology dominates the book, probably because one of the authors is a psychotherapist and one a psychiatrist. This perspective, together with the detailed interview materials, has produced a richly detailed picture of stepfamilies. I wonder, however, whether the immense amounts of survey data on these children and those like them could have been used *with* the interview material to paint a picture that was both detailed *and* panoramic.

*Placing Friendship in Context.* Edited by Rebecca G. Adams and Graham Allan. New York: Cambridge University Press, 1998. $54.95 (cloth); $19.95 (paper).

Maureen T. Hallinan
*University of Notre Dame*

Sociologists, more than most scholars, are expected to understand that context shapes human behavior. In the case of friendship, Adams and Allan's edited volume is a powerful reminder of this social fact. The book presents several studies that demonstrate the central role that context plays in the process of making, sustaining, and dissolving friendships. The authors argue that context not only provides opportunities for interaction and places constraints on behavior but also defines the meaning of friendship, its acceptable parameters, the nature of friendship interactions, and the consequences of friendship for individuals and the groups in which they are embedded. By highlighting the importance of context in shaping friendship patterns and dynamics, the book provides a salutary reminder that sociology has a unique perspective on this basic human relationship.

In their introductory chapter, Adams and Allan identify the various contexts in which friendships occur. The first context they point to is an individual's immediate environment or the circumstances of a person's life, such as family, work, and leisure activities. A second context is the social networks in which a person is involved, and which are both the product of social interactions and a determinant of interaction patterns. A third context is the community, which is characterized by economic, cultural, religious, and demographic factors. The broadest contextual level is a society whose economic and social structures affect the nature of personal relationships.

Given the editors' outline of relevant contextual influences on friendship, the reader might expect that the subsequent chapters would systematically link these contexts to friendship. The chapters are disappointing in this regard. Whereas each chapter makes a significant contribution to our understanding of friendship, few explain, in any detail, how specific contexts influence sociability. For example, in chapter 2, Mark presents research on the friendships of working-class women in an early 20th-century industrial plant. Although his study is intriguing, it leaves open the question of whether it is primarily societal norms, institutional practices, demographic characteristics of the community, network influences, or individual circumstances that determine the nature of these women's close friendships. Other chapters are equally vague in answering the question of how context affects sociability. While exhibiting insight and theoretical richness, most of the chapters lack specificity in linking contextual characteristics to friendship patterns. It may be that the lack of rigor in explaining how context affects social ties is a reflection of how difficult it is to define and bound social context and trace its effects.

To examine the effects of context on friendship systematically, one would need to do comparative analyses. The chapters in this text permit only the most general comparisons. Oliver's and Mark's chapters are comparable only in the sense that they are linked by the same time period, country, and concern about the effects of modernity on social relations. Oliver provides a theoretical analysis of the effects of industrial change on middle-class women's friendships in the United States, whereas Mark investigates the effects of industrialization on the friendships of working-class women in an early 20th-century America. The chapters suggest an influence of the industrial revolution on friendship patterns, but the class differences in the two studies, among other factors, preclude any general conclusions about societal effects on friendship.

The next three chapters are also weakly linked by a temporal and geographic commonality; they examine friendship patterns in middle and late 20th-century Britain. Allan looks at how social and economic contexts affect the formation and content of working-class male friendships, while Harrison examines the effects of friendships on a woman's perception of marriage and motherhood. O'Connor is concerned with how patriarchy and capitalism in a postmodern society influence women's friendships. Unfortunately, differences in class, gender, and context make most comparisons across these studies inappropriate.

Feld and Carter's chapter stands apart in its concentration on the effects of social networks on friendship. The authors argue that an important dimension of social context is the set of activities in which a person is involved. They claim that these activities influence the formation, endurance, and content of personal relationships and affect the density of an individual's social ties. Feld and Carter's chapter comes closest to specifying how contextual conditions influence individual behavior. A final substantive chapter by Adams examines the effects of technology on friendship by exploring how E-mail and the Internet affect the social

construction of personal relationships. This chapter, even more than the others, illustrates the difficulty in defining the context of social relations.

This book makes two significant contributions to our understanding of the nature of friendship. First, it reminds us that friendship is not only an individual matter but also a social event that occurs in a set of contexts that exert a powerful influence on its development. Second, the individual chapters are a rich source of theoretical ideas about the nature and content of sociability and provide a fertile background for future empirical studies linking context to friendship.

*The Environmental Crusaders: Confronting Disaster and Mobilizing Community*. By Penina Migdal Glazer and Myron Peretz Glazer. University Park: Penn State University Press, 1998. Pp. xxvi+218. $50.00 (cloth); $18.95 (paper).

Sherry Cable
*University of Tennessee—Knoxville*

This volume reports on Glazer and Glazer's research on community-based environmental activism in three countries in which they analyze the ways in which activists' backgrounds, motivations, tactics, and strategic alliances combine to contribute to the construction and maintenance of participatory, democratic society. Between 1991 and 1996, the authors conducted 140 interviews with environmental activists, public interest group members, journalists, union leaders, and corporate and government employees in the United States, Israel, and the former Czechoslovakia. Although their sample is not intended as representative of all environmental groups or activists in the three nations, it does reflect activism in response to the more prominent environmental issues in modern industrial societies: contamination from corporate and military production processes, potential contamination from proposed waste disposal facilities, and preservation of ecosystems and natural resources.

An introduction and the first chapter establish the context of democratic process and social capital in which the authors discuss grassroots environmental activists. Glazer and Glazer counter claims by social capital theorists of a recent decline in civil engagement in the United States that weakened democracy by short-circuiting the networks of mutual trust that facilitate citizens' collaborative efforts to solve common social problems. Instead, the authors suggest that citizens are joining a different type of organization. Rather than the traditional voluntary organizations, such as PTA and Rotary clubs, citizens are active in grassroots environmental organizations. These differ from traditional voluntary organizations; they challenge the status quo of business and corporate dominance at the community level and consequently revive or construct democratic institutions. Two major forces behind the recent emergence of these organizations are continuing discoveries of contaminated communities and

the end of the Cold War. Since 1978 and the revelations at Love Canal, communities are increasingly likely to confront health and safety threats from environmental contamination. The nature and extent of this contamination is more easily ascertained by citizens since the end of the Cold War, which significantly weakened the national security state and provided an opportunity to penetrate corporate and government secrecy and obtain information previously withheld from citizens.

Three subsequent chapters present cross-national examinations of grassroots environmental activists focused on specific topics: activists responding to threats linked to Cold War decisions, the role of mothers as organizers and leaders, and grassroots opposition to proposed garbage facilities. These chapters describe the personal transformation of core activists from uninvolved citizens to leaders in collective environmental efforts. Individual activists' profiles are provided, and similar stages are identified in their transformations: the initial, fearful awareness of a problem, the sense of betrayal by institutions they had trusted for protection, and collective action in which they create cultures of solidarity and build alternative power networks through strategic alliances. In the final chapter, the authors analyze the cultural, emotional, and social supports for activists' behavior, attributing their courageous responses to activists' reservoirs of social capital, their beliefs that they can make a difference, their abilities to accumulate evidence and expertise, and their faith in the justice of their cause.

Instead of presenting analysis, the book advocates; instead of offering insight, it incites. The book is neither theoretical nor even intensely sociological. Both the social movement and environmental sociology literatures suggest provocative research questions for such a data source. What was the role of recreancy, or institutional failure, in the personal transformation of activists? Did cases of *contamination* grievances differ from those of *proactive* grievances in the individual's path to activism? In the actions of elites? In the extent of intracommunity conflict? What was the political-economic context of the environmental issue? What payoffs accrued to whom? *Cui bono?* The cross-cultural nature of the data adds another dimension of frustration for analysts: the book lacks systematic cross-cultural comparisons. The researchers had three significantly different political settings for environmental activism: the United States, a modern capitalist state; Israel, a modernizing nation with continuing national security concerns; and the Czech Republic, a former Soviet ally left with a deteriorating transitional economy. Such data sets are seldom seen in studies of activism. Did paths to activism differ by country? Were organizational structures different? Did social control efforts vary?

Although the book may frustrate analysts, it is likely to inspire activists and other laypersons who are most likely to benefit from the book in practical ways. Such appears to be its audience, since theoretical references are contained in endnotes and the description of the methodology is in an appendix. The book is written without jargon and in a spirit of

equality and concern for the common good. Interested readers should include undergraduates in the social sciences and community organizers.

*Earthquake Fears, Predictions, and Preparations in Mid-America.* By John E. Farley. Carbondale: Southern Illinois University Press, 1998. Pp. xvi+198. $50.00 (cloth); $21.95 (paper).

Betty Hearn Morrow
*Florida International University*

While earthquake prediction, specific to time and place, has proven illusive to the scientific community, predictions by psychics or pseudoscientists occasionally make the news. Such was the case when self-proclaimed climatologist Iben Browning predicted that the New Madrid Seismic Zone, which runs through parts of Illinois, Missouri, Kentucky, Tennessee, and Arkansas, would experience a serious earthquake on December 2–3, 1990. This prediction received considerable media attention for several months leading up to the targeted date, in spite of its repudiation by most experts. Sociologist John Farley seized upon this event as a rare opportunity to measure attitudes such as risk perception and behavior, including preparedness, before and after an expected quake. With colleagues at Southern Illinois University at Edwardsville, Farley conducted four telephone surveys between October 1990 and May 1993, each in three different communities in southeast Missouri and southern Illinois. The results are reported in detail in this book and provide a "fascinating picture of changing beliefs and actions" (p. 5).

The first survey is the most unique since attitudes and plans were recorded prior to the predicted date. While many respondents were ambivalent, 53%–60% believed the occurrence of an earthquake early in December 1990 was somewhat or very likely. Most planned to take some kind of preparatory action, such as storing food and supplies. Between one-quarter and one-half planned to alter their schedule during that period. Those planning to leave the area varied from about 5% in St. Louis to over 20% in southeast Missouri. These numbers, while high, are in line with similar studies of "mass hysteria" events. Farley provides several explanations for the high acceptance rates, including the inexperience of Midwesterners in interpreting earthquake information and a lack of skepticism by the media.

When the same respondents were interviewed after the nonevent, fewer people reported actually altering their plans, not surprising given the expected desire for some to avoid ridicule. Nevertheless, about 15% of the total sample had kept their children home from school. This is not surprising considering some schools actually closed because of the prediction. While only 2% reported having stayed home from work, this involved about 200,000 people. Interestingly, the best predictor of sched-

ule changes was whether respondents perceived that their friends and neighbors believed the prediction. This supports the notion that risk perception is a social process, as well as the two-stage flow of communication model in which people confirm what they hear in the mass media through their social networks before deciding to take action.

The subsequent two surveys revealed that the heightened earthquake preparedness at the time of the prediction subsided somewhat but had sustained a plateau that was higher than before the Browning prediction. Farley uses this to illustrate the power of the mass media and proposes its more effective utilization by emergency planners. However, the false alarm did appear to reduce somewhat the number of people expecting a significant earthquake in the next 10–15 years. This research is of considerable practical importance since the New Madrid Seismic Zone is, in fact, at risk for earthquakes. According to scientists, there is a 30%–90% probability of a quake of the magnitude of 6.0 to 6.5 occurring within the next 50 years. Yet, prior to the Browning prediction, residents appeared to have little concern.

We need a better understanding of the factors associated with the perception of risk and, just as important, how risk perception then leads to concern and action. We are left with the uneasy notion that many, indeed most, ordinary citizens, and even some public institutions, distrusted the scientific community, or at least afforded enough credibility to Browning to hedge their bets. In truth, however, analysis of the data from this project explains very little of the variance in citizen attitudes and behavior. For example, the two attributes found to be significantly associated with belief in the prediction, education and gender (with college educated men being the least likely to believe the prediction), together only explained 8% of the variance.

Given the focus of my own work on the sociology of disaster, it is not surprising that I was interested in this work. However, most readers would not find the litany of research details particularly engaging. In reality, a shorter presentation of the findings, with the technical details confined to an appendix, would have resulted in a much more interesting presentation.

*Music and Social Movements: Mobilizing Traditions in the Twentieth Century.* By Ron Eyerman and Andrew Jamison. New York: Cambridge University Press, 1998. Pp. ix+191. $54.95 (cloth); $18.95 (paper).

Scott A. Hunt
*University of Kentucky*

Eyerman and Jamison conceptualize movements as "cognitive praxis." As knowledge-producing agents, movements reconstitute both cultures and politics by "providing a broader political and historical context for cultural expression, and offering, in turn, the resources of culture—tradi-

tions, music, artistic expression—to the action repertoires of political struggle" (p. 7). From this view, movements and popular music have a synergetic relationship. Movements provide conceptual space for the production of "truth bearing" music and charge it with "a special intensity and responsibility"; "truth bearing" music punctuates the vital essence of movement messages and suggests new lines for cognitive praxis (p. 77).

"Tradition" is key to understanding this relationship. Tradition includes both "regimens that delimit" and "resources" that enable action (p. 28). Eyerman and Jamison sketch several important traditions—populism, communism, as well as country, folk, blues, bebop, and gospel musics—and concentrate on pivotal artists, activists, and artifacts such as the Wobblies, Joe Hill, the *Little Red Songbook* (forever fanning the flames of discontent), John Lomax, A. P. Carter and the Carter Family, Carl Sandburg, the Seeger/Crawford Clan, the Weavers, Woody Guthrie, the Highlander Center, Lee Hays, Leadbelly, W. E. B. DuBois, and Langston Hughes.

Drawing upon the history of black music traditions, Eyerman and Jamison illustrate their thesis that music and movements exist in a synergetic relationship. Their analysis begins with DuBois. "For DuBois, the slave songs were not merely music, they were an expression of life rooted in a rural past, and more generally they reflected an experience all African-Americans could recognize. . . . Hearing them is hearing history; they are part of a collective memory that when recalled connects the present hearer with his past" (p. 75). Eyerman and Jamison's assert that DuBois's identification of black music (i.e., tradition building) contributed to the emergence of the Niagara Movement and the formation of the National Association for the Advancement of Colored People. Shortly after World War I, urban blacks—"New Negroes"—challenged DuBois's and other "Old Negroes'" conception of what constituted "authentic" black music. This challenge led to the building of the blues and jazz traditions. Eyerman and Jamison trace these traditions from the Harlem Renaissance of the 1920s, to the blues of Thomas Dorsey, to the bebop jazz of Charlie Parker, to the gospel of Mahalia Jackson, to the freedom songs of the 1950s, to the soul music associated with the black power movement, and finally to the hip-hop, disco, and rap musics of the 1980s and 1990s.

Following their analysis of 20th-century African-American music and movements, Eyerman and Jamison depict the synergetic relationship of music and politics in the 1960s. With notable variations by race, class, gender, and region, the sixties generation experienced a context of economic and political security, developed a unique generational self-awareness, and became enthusiastic consumers of mass media products, particularly phonograph records. The new generation's "mediability" led to a mixing of "high" and "low" culture. The politically conscious youth of the 1960s read philosophical and political treatises as well as considered serious themes presented in music.

The student movement, intimately connected to the Civil Rights movement, was concurrent with a folk music revival. For Eyerman and Jami-

son, Bob Dylan and Phil Ochs were the embodiments of the interplay between folk revival and student movement. The poetry and politics of both of these figures shaped and was shaped by the imagination of a generation in protest. Viewed together, the careers of Dylan and Ochs represent the eventual fate of the revival and movement. Dylan became a commercial success, a celebrity who no longer consistently carried movement ideals. In contrast, Ochs, untainted by commercialism and committed to progressive struggles, died young in a tragic, martyr-like suicide. With Ochs's suicide, the deaths of Janis Joplin and Jimi Hendrix, as well as increasing commercialism, "the intricate balance between culture and politics dissolved into its component parts, leaving both fundamentally different than before, but diffusing the revolutionary potential into disparate and often destructive directions" (p. 139).

Linking music and movements to the process of globalization, Eyerman and Jamison outline the impact the American experiences of the 1960s had on Sweden in the 1990s. Even though this case might be unfamiliar for many, Eyerman and Jamison's account of the synergetic relationship between music and politics in Sweden suggests the utility of understanding movements as cognitive praxis. Moreover, the chapter on Sweden shows how the cognitive praxis of music and movements is influenced by the folk traditions from many countries and is carried out for all the world to see via mass media.

*Music and Movements* is worthwhile reading for movement and popular culture scholars. I believe that this book could be supplemented with audio and video recordings as well as articles from popular print media (e.g., *Village Voice, Rolling Stone*) in undergraduate and graduate courses dealing with music, movements, and social change.

*Apparitions of the Madonna at Oliveto Citra: Local Visions and Cosmic Drama.* By Paolo Apolito. Translated by William A. Christian, Jr. University Park: University of Pennsylvania Press, 1998. Pp. xi+267. $49.50.

Sandra L. Zimdars-Swartz
*University of Kansas*

In May of 1985, in the small town of Oliveto Citra, two hours south of Naples, several young boys reported that they had seen the Madonna. Paolo Apolito, a professor of cultural anthropology at the University of Salerno, was present in the town for the first five days after this alleged sighting, recorded hundreds of conversations with skeptics and believers, and continued his investigations as the apparitions progressed in the years that followed. He has published two books on these apparitions, both concerned with analyzing how the subjective experiences of a group of children are transformed into socially constructed events that become a part of the cosmic drama of "Catholic visionary culture" (p. 15). *Apparitions of the Madonna at Oliveto Citra: Local Visions and Cosmic Drama*

is the second of the two books, published in Bologna in 1991 as *Il Cielo In Terra: Costruzioni Simboliche di un'Apparizione Mariana (Heaven on Earth: Symbolic Construction of a Marian Apparition)*. It is fortunate that William A. Christian, Jr., author of several important studies of Marian apparitions in Spain, recognized the significance of Apolito's contribution to the study of religious experience in general and apparitions in particular and undertook this English translation.

Apolito's premise is that what was universally described as a visual event (seeing the Madonna) was revealed through analysis to be an "event of words." In his initial study published in Bologna in 1990, *Dice che hanno visto La Madonna": Un caso di apparizioni in Campania (It is said that they have seen the Madonna: A case of apparitions in Campania)*, he argued that in the formative stage of the apparition, the children's subjective experiences became an "objective" event through a series of verbal transformations, moving from an indirect statement ("It is said they have seen the Madonna") to a direct declaration ("They have seen the Madonna"). This argument is briefly summarized in the prologue to *Apparitions of the Madonna at Oliveto Citra* and serves as an introduction to the apparitions and the analysis that follows.

In this second book, Apolito examines the negotiations of the various social actors as they endeavor to establish the symbolic world that surrounds the apparitions and provides the context for their meaning. In controlling and directing the meaning of the apparitions, Apolito gives pride of place to the parish priest and to the Queen of the Castle Committee he formed to conduct interviews with the seers and their supporters. Although hundreds of messages from the Madonna were reported, for example, the committee assumed the task of discerning the valid messages and published only a few of them. A "thinly disguised struggle" developed between the committee and the seers over the selection, and some seers even found their messages rewritten in words apparently deemed more appropriate. But the parish priest and his committee did not have absolute authority. Also important, for example, were the thousands of pilgrims who descended on the town. According to Apolito, most seers eventually had to accede to the expectations of believers that included, among other things, apocalyptic prophecies. The genius of Apolito's work is his ability to trace in detail the subtle verbal shifts and accommodations these and other constituencies made in their interactions with each other and to lay out in detail the process by which the context for the apparitions was created.

In the final chapter, Apolito considers the impact of technology on the worldview of apparitions. Cameras are ubiquitous at late 20th-century apparition sites, and the appearance of supposed miraculous images on Polaroid film has become an expected part of apparition culture for devotees. Apolito argues that one consequence of this proliferation of cameras has been to eliminate the privilege of seers since the miraculous image is now available to everyone. Moreover there is a new relationship to the divine since such things as eyeglasses, binoculars, and photography

"*capture* the image of the Madonna, as if they no longer let her offer herself to the chosen ones, but instead invade, like *paparazzi,* her 'private' realm" (pp. 207–8). Here the significance of the Italian title of the book becomes clear, for Apolito believes that this "vision by technology" has indeed succeeded in placing heaven on earth.

For those scholars who are convinced that there is no such thing as religious experience but only experiences that come to be understood as religious, *Apparitions of the Madonna at Oliveto Citra* is essential reading. Apolito convincingly demonstrates here a natural social process by which the exchange of words (broadly understood to include visual symbols) around the claim to a vision of the Madonna eventuates in a symbolic context within which meaning is assigned to the experiences of the seers and their followers. Although this symbolic system of meaning is but one such system in the lives of pilgrims, it is, nonetheless, a singularly important component in their formation of a multifaceted social identity.

*Schools of Asceticism: Ideology and Organization in Medieval Religious Communities.* By Lutz Kaelber. University Park: Pennsylvania State University Press, 1998. Pp. viii+278. $55.00 (cloth); $19.95 (paper).

Ilana F. Silber
*Hebrew University*

Lying at the core of Weber's classic thesis on *The Protestant Ethic and the Spirit of Capitalism* and still reverberating in his later comparative explorations of other religions was an intense concern with the uniquely "rationalizing" implications of inner-worldly asceticism as deployed in the specific context of Protestantism and, even more specifically, Calvinism. While Weber also explicitly addressed the need to search for the historical origins or antecedents of this distinctive form of religious asceticism in previous periods of Western history, he never actually carried out his plans for the kind of study that might have enabled him to examine preceding forms and phases of religious rationalization in the history of Western Christendom.

Searching for the medieval beginnings of inner-worldly ascetic rationalism, Lutz Kaelber's *Schools of Asceticism* sets out thus to correct what remained an important lacuna in Weber's own research program. Building upon a rich combination of primary and secondary sources, the outcome is a well-researched volume that not only demonstrates once again the ever surprising and enduring force of the Weberian research agenda but also stands out as a fine exemplar of the powers and significance of historical sociology more generally.

Kaelber asks to shift the attention away from orthodox medieval religion—be it in its clerical, monastic, mendicant, lay, or mystical trends— and to look among lay heterodox medieval religious movements. Specifically, he focuses on the early traces of what he calls (using a term perhaps

unduly loaded with anachronistic individualist connotations) the "em-
powerment of the self" associated with the capacity of rational inner-
worldly asceticism "to put all aspects of existence consistently and strin-
gently under ethical prerogatives, and to providing the means (given
certain institutional conditions) with which individuals could methodi-
cally permeate all spheres of society with ascetic practices" (p. 15). This
comes to correct what the author sees as two exaggerated tendencies in
extant sociological interpretations of medieval religion: the one privileg-
ing monasticism as the main early carrier of ascetic rationalism; the other
perceiving lay religion as only inclining to magical practices and thus
forming a largely obstructive, irrational force inimical to the development
of inner-worldly rational asceticism.

Of special value to Weber scholars perhaps is the careful rendering of
the development of Weber's views on medieval religion over the years
between *The Protestant Ethic and the Spirit of Capitalism* and later writ-
ings, often in close interaction with or response to the views of such major
contemporary figures as Troeltsch, Ritschl, or Harnack. And there is
much to commend and comment upon in Kaelber's own nuanced reas-
sessment of the legacy of medieval monasticism and of the impact of me-
dieval magic and sacramental practices—both backed by a useful and
critical reading of recent research on these topics.

On a more theoretical level, the argument seeks mainly to build upon
a shift of emphasis on Weber's part in his 1906 essay on "Churches and
Sects," moving away from the effects of religious ideas and their practi-
cal-psychological consequences, which were the central concern in *The
Protestant Ethic and the Spirit of Capitalism* and drawing greater atten-
tion to the influence of social organization—that is, the structure of and
interaction in sects and congregations—on patterns of social action. This
means approaching rational ascetic conduct as shaped not only by ideol-
ogy or "religious ethics" but also by social organization; or as Kaelber
sometimes rephrases it, to study the linkage between ideology and behav-
ior as mediated by specific organizational and historical contexts.

The strength of Kaelber's detailed study of two major heretical move-
ments in the High Middle Ages—Waldensianism and Catharism—does
lie indeed in its demonstration of the impact of organizational features:
that is, itinerant preachers (eventually acquiring quasi-priestly functions),
textual communities, congregations, and schools in the context of various
subvariants of Waldensianism; shops and workshops allowing for the
combined religious and vocational instruction of laity by virtuoso "per-
fecti" in the case of Catharism. These have clearly affected such impor-
tant variables as the channelling of asceticism into worldly professional
activities and the diffusion of heightened standards of lay, rather than
only virtuoso, methodical asceticism. However, it is not always too clear
whether what is demonstrated is the "mediating" impact of organization
upon ascetic conduct—the original theoretical intent—or more simply
and very convincingly (as in the case of Catharism most explicitly) upon
the eventual spread, success, and resilience of the movement. And it is

striking that the one and only movement that Kaelber describes as actually coming truly close to a pattern of inner-worldly asceticism, namely Austrian Waldensianism, displays the impact of doctrinal developments unique to this specific variant of Waldensianism—such as perseverance or proof of faith in conduct—and in this important case at least, perhaps more crucial than organizational structures.

More generally, this second part of the book displays a number of conceptual gaps, non sequiturs, and glissandos. While policies of active repression of heresies by ecclesiastical and secular authorities was the one institutional factor explicitly selected for systematic attention, the argument in fact rightly addresses a wider range of institutional and contextual factors that would have necessitated more systematic and explicit theorizing, such as socioeconomic status of followers, economic trends, urbanization, general religious climate, and so on. In this latter respect, in fact, one has to underscore the repeated reference to the widespread appeal of ideals of austerity and asceticism, in which one may perhaps detect a greater lingering effect of "orthodox" Christian and, even more specifically, monastic conceptions than favored by Kaelber's perhaps overly limited focus on heterodox religion and classical Weberian concern with "rationalization."

Only too infrequently indeed do sociologists have the scholarly skills and courage to enter the difficult and arcane domain of medieval history. As is made clear, however, from the addition of Kaelber's important work to previous efforts to either "reconstruct" Weber's conception or produce alternative interpretations of the medieval West, not only do such daunting enterprises add to our comprehension of the past, but they also provide us with precious avenues into such ever-perplexing sociological issues as the varying interplay of religious beliefs, behavior, and social structures, or the dynamics of continuity and change.

*Reluctant Modernity: The Institution of Art and Its Historical Forms.* By Alex Debeljak. Lanham, Md.: Rowman & Littlefield, 1998. Pp. xxiv+211. $55.00 (cloth); $22.95 (paper).

Albert J. Bergesen
*University of Arizona*

This book stops every few centuries from the Renaissance to the present and identifies what has been said about the institutional factors that are thought to be involved in art. There is a discussion of how the Renaissance supposedly brought art works out for consumption and onto the capitalist marketplace. Then, skipping a few centuries, a chapter on how the 18th century saw the emergence of mass publics and with that a sort of democratization of art—presumably in contrast to the Old Regime of more aristocratic art production/consumption, which lay between the Renaissance and the rise of "the bourgeois public sphere." Next, only a few

years skipped here, the early and mid-19th century and the assertion that Romanticism paved the way for the ideology of "art for art's sake." Then, on to a discussion of how "the historical avant-garde (Surrealism, Dadaism, Futurism, Constructivism)" challenged the earlier art for art's sake position. Finally, a discussion of "the postmodern art that emerges under the conditions of corporate capitalism in the mid-twentieth century . . . [and is] characterized by the fact that only half of the avant-garde program was actually realized" (p. xxii).

A lot of this assertion making is accomplished by reciting the opinions of others: "Benjamin famously defined . . .", "as Bourdieu would have it . . .", "in Bürger's view . . .", "although Kant argued . . .", "according to Weber's . . .", and so forth. This brings us to the heart of the matter. What is to be said of these various assertions? Its hard to say, in that they are mostly just put out there, and often in the form of "as Bourdieu would have it." After a while you wish the author had a distinctive position rather than repeating, rehashing, and glossing over other cultural commentators on the history of art. The author is a poet and cultural critic, so the absence of a distinctly sociological analysis is not all that surprising, although the cultural commentary could have had more original claims, or evidence, or just illustrations so that the reader would have an idea of whether "in Bürger's view" seemed supported by the historical record.

In the author's defense, some institutional areas seem hard to bring under systematic sociological analysis. A sociology of society seems easiest. We define society, and we study it. When we get close to the economy, though, we seem to fall prey to economics and lose our nerve to theorize in purely sociological terms. The sociology of law is, by and large, a weak area in good part because there is too much jurisprudence and not enough sociology. There is probably something like a general principle here: the more an institution has a theory of itself, the more our conception of its structure and process is alloyed with its native accounts. The sociology of art seems like this too. Accounts by art historians, critics, art theorists, and generalized cultural commentators always seem to get mixed up with our sociological understanding of what is happening. Is, for instance, the argument that, "modernist art embodied a willful retreat from the repulsive machinations of a capitalist world" (p. 139) the result of an analysis of the real art world, or of what some artists said about what they were doing, or conventional modernist ideology? I am not sure. Some artists no doubt felt this way, some art commentators no doubt argued this, but a lot of artists had dealers, worked with galleries, and sold their paintings on the capitalist market. The old heroic story of modern art, repeated, and worse repeated by retelling the positions of other writers, leaves a lot to be desired in a book on the institutions of art.

If there is a singular shortcoming of this book, then, it is that it just repeats commonplace arguments about art movements from art history: the Renaissance took art from religion and brought it into the marketplace; the avant-garde was compelled to transform the institutions in

which art was produced; capitalism distorts artistic production; and so forth. There is little new here that has not already been argued by the many authors frequently quoted, and there are few examples and little evidence that any of these claims hold up. So, it is hard to evaluate. A good sociology of art, though, probably needs to do more than repeat the traditional stories about the history of Western art.

*The Stars Are Not Enough: Scientists—Their Passions and Professions.* By Joseph C. Hermanowicz. Chicago: University of Chicago Press, 1998. Pp. xv+268. $45.00 (cloth); $15.00 (paper).

Theodore P. Gerber
*University of Arizona*

How do scientists cope with the inevitable recognition that they will not become the next Albert Einstein? How do they redefine their career ambitions in reaction to exigencies imposed by their stage in the life course and the institutional setting in which they work? Joseph C. Hermanowicz addresses these questions based on 60 in-depth interviews with physicists at six different universities. He focuses on scientists' professional self-identities—the "narratives" they construct about their position within the moral order of their profession, the meaning of their work, and how it shapes their broader conceptions of self. The emphasis on how scientists perceive their careers sets this book apart from traditional studies of stratification within science and from constructivist analyses of how interested agents create scientific knowledge.

Hermanowicz proposes that scientists work in one of three "academic worlds," each with distinct opportunities, incentives, and socially enforced norms. These worlds shape professional self-identities so decisively that "clones at the starting gate in each world would very likely look altogether different by the time they reach the finish line of their careers" (p. 153). *Elite* departments provide ample resources and rewards for excellence in research, the sole criterion for evaluating professional and personal worth. Elite scientists hold themselves and each other to uncompromising standards, severely sanctioning those who do not continually produce acclaimed research. Success only ratchets expectations higher, and one derives status solely from one's most recent work. Thus, perhaps surprisingly, although they are most successful in terms of research, elites fear professional failure more than other scientists do. They also suffer more from this fear, because science looms large in their overall self-identities.

At the other extreme, *communitarian* departments downplay research in favor of teaching and administration. Whereas elite institutions nurture ambition to a fault, meager local incentives and resources compel communitarians to abandon dreams of making their mark through re-

596

search accomplishments. They perforce embrace alternative career goals such as success in teaching. They value interpersonal aspects of work and find fulfillment in pursuits outside of science. While elite "moral careers" progress through a series of well-defined stages, communitarian careers take diverse paths, since communitarians accept a variety of career aims as legitimate. Science figures less centrally in how communitarians view themselves and their colleagues. Although they may lament the state of their family situation or tennis game, communitarians voice few regrets relating to their modest research accomplishments.

*Pluralist* departments combine both elite and communitarian elements. This points to a problem with Hermanowicz's schema: distinctions among departments are surely less clear-cut than a three-category typology implies. Within every department—perhaps within the self-identity of every scientist—elite and communitarian narratives coexist, if not always harmoniously. Rather than falling into one of three distinct science worlds, departments and scientists probably range along a continuum from the more elite to the more communitarian. Implicitly recognizing this, Hermanowicz does not specify criteria for classifying departments and provides examples of communitarian sentiments among elite informants, elite aspirations among pluralists, and so on.

The book's strength lies in the interview materials, which render professional identities effectively and often poignantly. Consider the vivid intensity of ambition of "Silverman," whose "dream is to discover some fantastic new effect that knocks the socks of my friends and colleagues, that knocks the socks off the community, so that when I walk down the corridor, the young students know me and say, 'There goes [Silverman], he invented the [Silverman] effect.' That's what I want; I want my effect. . . . I can even smell what it's like already" (p. 110). Many passages where informants recount how their aspirations have evolved, or assess their achievements and failures, will resonate with readers who themselves work in academia.

Although the interviews offer fascinating reading, the analysis disappoints. We already know that age and academic position influence scientific productivity. The social-psychological consequences of these effects could provide insights into the life course, careers, and identity construction, but the book lacks the conceptual framework needed to draw general conclusions. Hermanowicz describes the contrasts among his three worlds, but he never methodically explains how they are created, sustained, and developed over time. Exactly how, for example, do elite scientists exert the mutual social control that keeps standards so high? Hermanowicz wavers on certain key issues, such as whether or not communitarians also aspire to greatness early in their careers. Others, such as how graduate school socialization or the academic labor market shape aspirations, are touched on, but require more sustained attention. I found the self-described "inductive" presentation of the interview materials distracting, the chapter organization inconsistently followed. Suggestive

hints of explanations abound, at times in the form of poetic allusions. But hints are not enough. We need more consistent and cohesive analysis to guide us to the theoretical punch lines.

*Claiming Power in Doctor-Patient Talk.* By Nancy Ainsworth-Vaughn. New York: Oxford University Press, 1998. Pp. xii+212. $45.00 (cloth); $19.95 (paper).

Douglas W. Maynard
*Indiana University.*

The agenda is clear in this book. Nancy Ainsworth-Vaughn believes that medical care can and should be a partnership between physician and patient. Her research on this relationship shows both that at times it is not such a partnership and that at times it is. Making suggestions about how to enhance partnership, Ainsworth-Vaughn does not make idealistic or philosophical arguments but ones grounded in her empirical approach. A major contribution of this book is in countering the usual one-sided view of how dominant and powerful physicians are in the medical encounter. Where patients literally have been *lost* in some accounts of the doctor-patient relationship, Ainsworth-Vaughn *finds* the patient again.

Chapters cover topic control (and gender), question-answer sequences, the use of rhetorical questions, storytelling, and diagnosis. At the risk of simplification, findings can be stated in general terms. In considering gender and topic control, Ainsworth-Vaughn analyzed topic transitions in 12 encounters involving eight physicians (four male/four female) and eight patients (all women). She developed a way of coding topic transitions according to whether they were a reciprocal decision or unilaterally imposed, and computed ratios of the former to the latter. Patients had a ratio of reciprocal to unilateral topic transitions of 13.5 to one, while for physicians it was only 2.5 to one. This physician ratio was further decomposed into 5 to one for female doctors and 1.4 to one for males. In these terms, physicians dominate and control topics in the encounter, and male physicians do so more than female physicians. (There were no male patients and consequently no topic changing ratios to report for this class of participant.)

Investigating the distribution of questions, Ainsworth-Vaughn has surprising findings. In two of the best-known previous studies, Richard Frankel found patients initiating questions only 1% of the time, and Candace West found a 9% question-producing rate for patients, but Ainsworth-Vaughn turns up 38.7% of questions being asked by patients and only 61.3% being asked by physicians. She considers various explanations for this anomaly, including the type of data she has (oncology, wherein patients can be expected to ask more questions than in more regular visits). However, Ainsworth-Vaughn's primary suggestion is that her research simply contradicts the stereotype of the passive patient.

Consideration of the near 60/40 balance between physicians and patients in asking questions gives way in a qualitative chapter on rhetorical questions to further puncturing of the "dominating physician" notion. Ainsworth-Vaughn shows that when physicians use such questions, it can be a way of exhibiting their structural power. Nevertheless, rhetorical questions may build in ambiguity that mitigates this power and is at least a ritual nod of the head to patients' autonomy. Furthermore, if the medical encounter is "fundamentally asymmetrical" in terms of physicians' holding structural power, patients themselves nevertheless know how to exploit rhetorical questions and ambiguity to achieve their own ends. They are not innocent or hapless bystanders to physician control.

And finally, Ainsworth-Vaughn considers how diagnosis occurs in the medical encounter. Recognizing that medical encounters are inherently asymmetrical because of physicians' expert knowledge and legitimated authority, she proposes that mutual storytelling is somewhat able to neutralize this asymmetry. That is, if doctors listen to patients' explanations for their disease, and allow them to present sometimes anomalous-seeming narratives, there is reciprocal determination of diagnostic outcomes. At the very least, physicians may hear something that leads them on a different path than their otherwise one-sided considerations might do.

The sociologist will have trouble with some conceptual matters in this book. For instance, where we usually consider *power* to involve a relationship, Ainsworth-Vaughn defines it as "implementing one's agenda." Hence, independent of whether the patient resists or cooperates, being able to ask questions is an exercise in power. Given the centrality of this concept to the book, and although several pages are devoted to different aspects, I would have preferred more extensive discussion of power and what it means to have it in the context of the medical encounter. On the other hand, a mixture of methods will appeal to sociologists who prize triangulated research. Ainsworth-Vaughn conducts counting and distributional studies in addition to doing ethnographic discourse analysis, which is her main approach. She argues, against conversation analysts for instance, that it is necessary to immerse oneself in the sociocultural context to fully analyze utterances and interaction. Indeed, her interviews with participants are a nice complement to her recordings, but I (as one who does conversation analysis) was not fully convinced of the methodological point.

Ainsworth-Vaughn began by recording her own encounters in oncology, and her cancer diagnosis and experience unobtrusively inform the study. However, recordings from her examinations do not become part of the analytic chapters in the book. Rather, she obtained permission to audiotape in a private-practice oncology setting and gathered data for seven years, including 40 patients, 13 physicians, and following some doctor-patient pairs over time. She conducted field observations and numerous interviews as well. In this impressively hard-wrought effort, Ainsworth-Vaughn is alive to the subtlety and nuance of medical encounters and refuses to see them in simplistic terms. Again and again, she

stresses that no single stroke will work to paint a picture of this encounter. Her multicolored venture is a welcome change from the usual drab canvas. The final chapter, "Active Patients and Cooperative Physicians" draws implications both for patients and for physicians who are interested in the partnership idea.

*Sex and Medicine: Gender, Power and Authority in the Medical Profession.* By Rosemary Pringle. New York: Cambridge University Press, 1998. Pp. x+240. $64.95 (cloth); $22.95 (paper).

*The Woman in the Surgeon's Body.* By Joan Cassell. Cambridge, Mass.: Harvard University Press, 1998. Pp. vii+267.

Christine Williams
*University of Texas, Austin*

When occupations change their gender composition, and go from being predominately male to predominately female, typically two things happen: the status and pay of the occupations decline and the culture of the occupations becomes feminized. Is this what is happening in medicine? An unprecedented number of women are now doctors, and in many medical schools, they now make up half of the student body. Does this mean that medicine will lose its high rank among the professions? Are women transforming medicine into a "feminine" occupation and in the process making it more humane and democratic?

These questions are answered in two new books on women in medicine. *Sex and Medicine,* by Australian sociologist Rosemary Pringle, examines the status of women throughout the medical profession in Britain and Australia. Joan Cassell, an American anthropologist, focuses on women surgeons in the United States and Canada in her book, *The Woman in the Surgeon's Body.* Pringle's book is based primarily on in-depth interviews conducted with 150 women doctors, while Cassell examines the lives of 33 women surgeons using a combination of ethnographic and interviewing techniques. Despite their geographical and methodological differences, the two books reach remarkably similar conclusions. Women continue to be marginalized within the profession, but there are signs that they may be changing the profession for the better.

Surgery is the specialty that has been most resistant to women, and it is there that the barriers to women are most apparent. Cassell gives many blood-curdling examples of sexism, such as the Monday morning staff meetings when each (male) surgical resident was required to give a nonverbal signal to indicate "what kind" of sex he had on the weekend, while the lone woman sat there passively. Like a number of men in many male-only occupations, surgeons often define their masculinity through their work and view women as devaluing and threatening the worth of their

achievements. They express their misogyny through such sexist rituals intended to ostracize women.

The gendered organizational structure of surgery forms another formidable obstacle to women. To put it bluntly, women with children are not welcome and need not apply. It is virtually impossible to complete surgical training and to embark on a successful career, and also bear and raise children. Compared to North America, there are more opportunities in Britain and Australia for part-time residencies and leaves-of-absence from medical school, both of which enable some women to combine childbearing with medical careers. But even so, Pringle points out that those women who avail themselves of these options are considered unreliable and not devoted enough to their careers, and they tend to be disadvantaged in a competitive job market.

To explain the persistence of gender discrimination in medicine, both Cassell and Pringle rely extensively on Pierre Bourdieu's theory of habitus. They are drawn to this perspective because it grapples with the embodied aspect of social inequalities. Women who grow up in a sexist society literally incorporate gender into their ways of being: they talk, think, interact, walk, and generally inhabit their bodies in distinctively female ways. This female "habitus" acts as a barrier to their integration into the masculine culture of medicine, but it also brings with it the hope for a more humane medical practice.

Both books contribute to the feminist project to understand occupational segregation and workplace discrimination, but oddly, both authors express ambivalence toward feminism. Pringle notes that feminists active in the women's health care movement have treated doctors as enemies of women for their attempts to control and pathologize women's bodies. This puts women doctors in a peculiar position vis-à-vis feminism. Although they are clearly the beneficiaries of the feminist movement, they nevertheless embrace many of the so-called elitist and patriarchal values attributed to the medical profession by feminist health activists. Consequently, it is not surprising that few of the women studied in the two books embraced feminism. But it is surprising that Pringle and Cassell also distanced themselves from the label.

Just as there is substantial overlap in the findings of these two books, there is also overlap in their omissions. Missing from both is an analysis of doctors' location in the class structure. Neither author emphasizes the fact that the women they studied are *wealthy*. We are not given income data on the samples. We catch only glimpses of their current class position in Cassell's book, when she describes the doctors' lovely homes, diamond earrings, or luxury cars. This is an important omission, because it helps to explain how women benefit from the current organization of medicine, despite its sexist flaws. Understanding the benefits as well as the drawbacks of the current system could also help to explain women's resistence to feminism.

Also missing in both books are the voices of men. The sexism of the

medical profession is explored entirely from the vantage point of women. While there is no reason to challenge the truthfulness of the women's claims, the arguments could have been more compelling if the authors had chosen to include men's perspectives in their books. Both could have easily done this: Pringle conducted 30 interviews with men doctors as part of her study, and Cassell wrote a previous book on male surgeons; however, neither draw upon these data to examine how men have reacted to the inclusion of women. Including men's opinions is important not because it would "balance" the books. Rather, I wanted to hear if the women were having any positive impact on men doctors. If, as both authors argue, women have the capacity to change medicine in a more humane direction, then changing men must be part of the process. Is there any evidence that men are learning from the women?

Although I have stressed their commonalities, there are important differences in the two books, especially with regard to their organization and tone. *Sex and Medicine* is organized by medical speciality, with chapters devoted to gynecology, surgery, internal medicine, anaesthesia, psychiatry, and general practice. (Other chapters address women doctors' relationships with nurses, and the women's health movement.) This is an unusual choice granted Pringle's stated interest in challenging occupational boundaries within medicine as part of the move to democratize the profession. These chapters follow a similar format, sketching the history of each specialty, followed by a history of women's participation in the specialty, and then a discussion of the interview findings regarding the barriers and opportunities to women. Similar themes crop up in each chapter, making the book seem overly repetitive.

*The Women in the Surgeon's Body,* in contrast, reads like a novel. Cassell uses a much smaller sample, with whom she spent a longer period of time collecting information. Each woman in the study agreed to be "shadowed" by Cassell for two to five days and then interviewed over dinner. Her analysis is presented through stories, many taken directly from her field notes. She conveys a very physical sense of women's bedside manner, how they comport themselves with colleagues and interact with patients and nurses—and researchers. Cassell is very much present in her book; readers get to know her as intimately as the women doctors she studied.

But the similarities of the two books overwhelm their differences. These books describe a profession in the midst of a radical social change. The picture they paint of women in medicine is a mixed one, with vestiges of extreme sexism and hints of humanistic possibilities. Sociologists of work, gender, and medicine will appreciate these updates on this revolution in progress.

# Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods[1]

Robert J. Sampson
*University of Chicago and American Bar Foundation*

Stephen W. Raudenbush
*University of Michigan*

This article assesses the sources and consequences of public disorder. Based on the videotaping and systematic rating of more than 23,000 street segments in Chicago, highly reliable scales of social and physical disorder for 196 neighborhoods are constructed. Census data, police records, and an independent survey of more than 3,500 residents are then integrated to test a theory of collective efficacy and structural constraints. Defined as cohesion among residents combined with shared expectations for the social control of public space, collective efficacy explains lower rates of crime and observed disorder after controlling neighborhood structural characteristics. Collective efficacy is also linked to lower rates of violent crime after accounting for disorder and the reciprocal effects of violence. Contrary to the "broken windows" theory, however, the relationship between public disorder and crime is spurious except perhaps for robbery.

> The answer to the question of how city life was to be possible, then, is this. *City life was made possible by an "ordering" of the urban populace in terms of appearance and spatial location such that those within the city could know a great deal about one another by simply looking.*
> (Lyn Lofland, *A World of Strangers: Order and Action in Urban Public Space,* 1973, p. 22; emphasis in original)

Visual signs of social and physical disorder in public spaces reflect powerfully on our inferences about urban communities. By social disorder, we refer to behavior usually involving strangers and considered threatening,

such as verbal harassment on the street, open solicitation for prostitution, public intoxication, and rowdy groups of young males in public. By physical disorder, we refer to the deterioration of urban landscapes, for example, graffiti on buildings, abandoned cars, broken windows, and garbage in the streets. Visible evidence of disorder, or what Albert Hunter (1985) calls "incivilities," have long been noted as central to a neighborhood's public presentation (Goffman 1963). Jane Jacobs's classic observation of urban life in the 1950s even then evoked a concern with the threats of disorder to neighborhood civility (1961, pp. 29–54), especially the negotiation of public encounters in the "world of strangers" (Lofland 1973).[2]

The streets, parks, and sidewalks still belong to no one and therefore to everyone. Disorder continues to be of theoretical interest precisely because of its visual salience and symbolism regarding the use of such spaces. Even if we wish it were not so, disorder triggers attributions and predictions in the minds of insiders and outsiders alike. It changes the calculus of prospective home buyers, real estate agents, insurance agents, and investors and shapes the perceptions of residents who might be considering moving. Evidence of disorder also gives a running account of the effectiveness of residents seeking neighborhood improvement, and that record may encourage or discourage future activism. Physical and social disorder in public spaces are thus fundamental to a general understanding of urban neighborhoods.

Neighborhood disorder has more specific bearing on the study of crime as well. Research has established connections between disorder and both fear of crime and crime rates (Skogan 1990; Kelling and Coles 1996). In fact, a reigning theory posits that minor disorder is a direct cause of serious crime. Originators of the "broken windows" thesis, Wilson and Kelling (1982) argued that public incivilities—even if relatively minor as in the case of broken windows, drinking in the street, and graffiti—attract predatory crime because potential offenders assume from them that residents are indifferent to what goes on in their neighborhood. The metaphor of

[2] Goffman (1963) goes back yet further, to the obligation in medieval times to keep one's pigs out of the streets. In this case, the norms regulating public order refer not just to face-to-face interaction among strangers or acquaintances, but the visual ordering of the physical landscape (1963, p. 9). For example, Goffman writes of expectations regarding the maintenance of sidewalks and keeping the streets free of refuse Hence, disorder in public places may be conceived in physical as well as social terms

broken windows is apt, insofar as the theory asserts that physical signs of disorder serve as a signal of the unwillingness of residents to confront strangers, intervene in a crime, or call the police (Greenberg and Rohe 1986; Skogan 1990, p. 75). Proponents thus assume that both physical disorder and social disorder provide important environmental cues that entice potential predators. The "broken windows" thesis has gained ascendancy in criminology and has greatly influenced public policy, leading to police crackdowns in numerous cities on the manifestations of social and physical disorder. New York City is the best-known example of aggressive police tactics to control public incivilities (Kelling and Coles 1996, pp. 108–56; Kelling 1998).

Taking seriously the idea that visual cues matter, this article applies the method of systematic social observation (SSO) to the study of social and physical disorder in urban neighborhoods. We depart from prior research in three ways. First, we describe novel systematic procedures for collecting observational assessments of public spaces using videotaping procedures that produce a permanent visual record amenable to later coding and reinterpretation based on emergent insights.[3] Second, we formulate a hierarchical item-response model that identifies sources of error in aggregating across observed disorder items within block faces and in aggregating across block faces within some 200 census tracts. The method yields high tract-level reliabilities for assessing both social disorder and physical disorder.

The third and major goal of the article is to assess the sources and consequences of neighborhood disorder. We do so by testing the association of systematically observed disorder with independent measures of officially recorded and survey-reported crime, census-based sociodemographic composition, and a survey-based measure that taps the collective efficacy of residents in achieving informal social control. A theory combining structural constraints with local collective efficacy is presented as an alternative to the "broken windows" interpretation of the disorder-crime link. We also assess broader implications for reinvigorating the study of urban communities based on systematic observation and video-based approaches.

## SYSTEMATIC SOCIAL OBSERVATION OF DISORDER

In the spirit of the early Chicago school of urban sociology, we believe that direct observation is fundamental to the advancement of knowledge

[3] Building on Lofland (1973, p 19), we define public spaces as those areas of a neighborhood to which persons have legal access and can visually observe—its streets and sidewalks, its parks, its places of public accommodation, its public buildings, and the public sectors of its private buildings.

(Park and Burgess 1921; see also Whyte 1988). As Andrew Abbott (1997) notes, one of the hallmarks of the Chicago school was its concern with observing public places—not just abstract variables, but the sights, sounds, and feel of the streets. Attempting to systematize such approaches, more than 25 years ago Albert J. Reiss, Jr. (1971) advocated systematic social observation as a key measurement strategy for natural social phenomena. By systematic, Reiss meant that observation and recording are done according to explicit rules that permit replication; he also argued that the means of observation, whether a person or technology, must be independent of that which is observed. By natural social phenomena, Reiss (1971, p. 4) meant "events and their consequences, including properties of organization, can be observed more or less as they occur." As his main example, Reiss described systematic observations of police-citizen encounters. He also noted the general import of the SSO method for assessing physical conditions and social interactions within neighborhood settings that survey respondents may be incapable of describing accurately.

Despite the potential yield of direct observation, the majority of research studies linking signs of disorder with fear of crime and criminal victimization have been based on residents' subjective perceptions drawn from survey responses.[4] The typical strategy in survey research has been to ask residents how much of a problem they perceive disorder to be; the standard finding is that perceptions of disorder predict fear of crime (Skogan 1990; Perkins and Taylor 1996; Taylor 1997, 1999). The dearth of independent assessments of neighborhood disorder poses a special problem for interpreting this linkage. As Taylor (1999) has argued, the high correlation between fear and disorder may arise in part from shared survey-method variance. More fundamentally, however, the perception of disorder seems also to reflect a psychological construct—perhaps fear itself (Garofalo and Laub 1978; Rountree and Land 1996). Residents fearful of crime report more disorder than do residents who experience less fear, even though both sets of observers are reporting on the *same* neighborhood (Taylor 1999; Perkins, Meeks, and Taylor 1992). In this scenario, the fear (or vulnerability) of residents might be said to induce perceptions of disorder. Even the disorder-crime link is problematic, since victimization experiences are usually measured in the same surveys used to assess (perceived) disorder.

One of the primary obstacles to bringing independent and systematic social observation to bear on this conundrum has been methodological uncertainty, not just on how to properly conduct such observations, but on how to properly assess their measurement properties at the neighbor-

---

[4] Useful reviews of the empirical literature are found in Perkins and Taylor (1996), Taylor (1997, 1999), Skogan (1990), and Skogan and Maxfield (1981)

hood level (Raudenbush and Sampson 1999*b*). Another concern has been cost, even though direct observations are potentially less expensive than household surveys, with listing, screening, broken appointments, and response rates eliminated. Yet another obstacle has been conceptual, stemming from underappreciation of the yield of systematic observation for one of the fundamental cleavages in sociological criminology—the reliable and valid measurement of crime and deviance. Perhaps most important, however, has been the psychological reductionism that flows from the dominant theoretical and empirical focus on individuals.

An exception to the lack of independent observations of disorder at the level of ecological units rather than persons is found in research by Taylor and colleagues (Taylor, Shumaker, and Gottfredson 1985; Taylor, Gottfredson, and Brower 1984; Covington and Taylor 1991; Perkins et al. 1992; Perkins and Taylor 1996; see also Mazerolle, Kadleck, and Roehl 1998).[5] Using observations conducted by teams of trained raters walking the streets, Taylor et al. (1985) assessed 20% of all occupied face blocks in 66 Baltimore neighborhoods. A face block is the block segment on one side of a street. They identified two physical dimensions that stood out empirically: physical incivilities and nonresidential land use. These two dimensions were reliable in terms of individual-level psychometrics (e.g., Cronbach's alpha; interrater reliability) and were related as expected to perceived disorder and fear of crime derived from neighborhood surveys. More recently, Perkins et al. (1992) examined on-site assessments of block-level physical incivilities in Baltimore. Controlling for social factors, physical incivilities predicted perceptions of crime-related problems. Yet, using similar procedures in a different city, Perkins et al. (1993) report that residents' perceptions and an independent rating of physical disorder were not significantly correlated. Observed environmental items correlated more strongly with multiple indicators of subsequent block crime than did residents' perceptions of the environment. Interestingly, residents' perceptions of physical disorder correlated positively with fear, but not after controls were introduced for income, stability, and racial composition.

Overall, then, the research record is mixed and curiously imbalanced. Although specified as an ecological construct, neighborhood disorder has been investigated mainly using individual perceptions and individual-level research designs. The number of studies employing observational ratings across multiple ecological contexts is small, and the correlation of observed disorder with subjective perceptions varies by level of aggrega-

---

[5] An early version of systematic observation based on single interviewer ratings in a neighborhood survey was used in Taub, Taylor, and Dunham (1984) and Skogan (1990).

tion, type of measure, and study site. We therefore approach the study of disorder from an integrated observational, survey, and record-based approach at the neighborhood level. Just as important, we offer a theoretical framework on the sources and consequences of public disorder that challenges the prevailing view on broken windows and crime.


Rethinking Disorder

Rather than conceive of disorder as a direct cause of crime, we view many elements of disorder as part and parcel of crime itself. Consider the typical items used to define social disorder, such as solicitation for prostitution, loitering, and public use of alcohol or drugs. Consider also physical "incivilities," such as graffiti, smashed windows, and drug vials in the streets. All of these are evidence either of crimes (whether concurrent with the observation, as in drug use, or physical evidence of recent acts, as in drug vials on the sidewalk) or ordinance violations. Although ordinance violations like drinking in public and many "soft crimes" like graffiti may not be judged as particularly serious, this is an evaluation or classification issue and not a statement on etiology. As Gottfredson and Hirschi (1990, pp. 42–43) have argued, criminologists often mistake differences in crime seriousness (classification) for differences in causal mechanisms. But as the long history of developmental research on juvenile delinquency has instructively shown, minor offenses usually do the best job of discriminating individual differences in later serious crime. In fact, early smoking and truancy are among the most reliably measured predictors of the propensity to serious adolescent delinquency (Farrington 1979), presumably because these acts all have the same antecedents (Gottfredson and Hirschi 1990). Relatedly, Hagan and McCarthy's (1997) recent ethnographic and quantitative study of two Canadian cities demonstrates the close connection of predatory youth crime to street life and settings of public disorder (e.g., prostitution, vagrancy, drug selling).

Applying the logic of Gottfredson and Hirschi (1990) to the present neighborhood-level case, a reasonable hypothesis is that public disorder and predatory crimes are manifestations of the same explanatory process, albeit at different ends of a "seriousness" continuum. Even those elements of disorder not obviously criminal in nature (e.g., garbage, vacant housing) are either violations of an ordinance (as in littering, slumlord abandonment) or may be conceptualized as sharing a similar causal structure and thus predicted by similar mechanisms (Hunter 1985). Concretely, for example, it does not seem to us persuasive to argue that graffiti causes robbery. Lack of social control might cause both graffiti and robbery; if so, one should measure the specified causal mechanism rather than (tautologically) inferring lack of order from graffiti and then using it to explain

robbery. What makes this conceptual move significant, in our view, is that it provides the opportunity to *observe* and hence systematically measure important manifestations of crime-related processes. Muggings, assaults, and rapes might be impossible to reliably observe, but vandalism, prostitution, gang congregation, and evidence of drug use can, in principle, be observed by all, whether residents, business people, visitors, possible investors, local activists, or potential offenders. Sociologists of crime have debated for at least 30 years the relative merits of survey-reported crime (whether offending or victimization) relative to official police records (Short and Nye 1957). By recasting disorder in the theoretical terms of crime, an observational window is opened on a new alternative for testing neighborhood-level theory.

Of course, not all environmental observations tap disorder. Researchers have profitably examined land use (e.g., mixed residential-business), the presence of bars, street layout, traffic patterns, and housing structure, all of which are conceptually distinct from crime (e.g., Perkins and Taylor 1996; Taylor 1999; Perkins et al. 1992). Such environmental features are legitimate targets of observation, and we too incorporate them in our methodology. Moreover, our argument is not that social and physical disorder are unimportant for explaining neighborhood dynamics. To the contrary, our framework suggests that while both crime and disorder reflect common origins, crime may be less relevant for understanding processes such as population abandonment and the perceived incivility of urban life because it is largely unobserved (see also Jacobs 1961; Skogan 1990). Corresponding to the "text" from which all key actors in a neighborhood read, we propose that disorder is the more visually proximate or immediate neighborhood cue of theoretical interest, even if it is not a direct cause of further crime.

## STRUCTURAL CONSTRAINTS AND THE AGENCY OF SOCIAL CONTROL

The theoretical framework that guides our assessment of disorder stems from a balancing of structural constraints with recognition of purposive social action. In the study of crime and disorder, attention has focused primarily on structural dimensions of an economic nature over which residents are thought to have little control—especially concentrated poverty and its associated lack of social resources. The ecological concentration of disadvantage means that the poorest neighborhoods tend to have not only the lowest incomes but also higher rates of unemployment, financial dependence, and institutional disinvestment (Wilson 1987; Land, McCall, and Cohen 1990; Hagan and Peterson 1995). Economic deprivation is relevant in that repairing buildings and cleaning up residential and commer-

cial areas requires money. Because areas of concentrated disadvantage find it difficult to support viable commercial enterprise, many stores and apartments will be vacant, giving little incentive for investors to repair their properties. The density of children in single-parent families, which is ecologically concentrated with poverty and resource dependence in U.S. cities (Land et al. 1990), adds another layer of difficulty to the always challenging task of supervising children and adolescent peer groups.

Structural constraints are not just economic in nature. Systemic theories of urban communities (Kasarda and Janowitz 1974) have long pointed to the importance of residential stability as a major feature of urban social organization. High levels of home ownership and low transience work together to instill in residents a "stake in conformity," in this case to neighborhood well-being. The formation of social networks that undergird local ties and attachment to place is also linked to residential stability (Sampson 1988; Taylor 1997). By stability we do not mean lack of change but rather the social reproduction of neighborhood residential structure, typically when population gains offset losses and home values appreciate.

Against the backdrop of resources and stability, a number of other structural constraints impinge on the ability of neighborhoods to counteract public incivilities, including the sheer density of population, nonresidential land use, public transportation nodes, and large flows of population that overwhelm local services. The "routine activities" perspective (Cohen and Felson 1979) builds on the insight that predatory crime involves the intersection in time and space of motivated offenders, suitable targets, and the absence of capable guardians. Because illegal activities feed on the spatial and temporal structure of routine legal activities (e.g., transportation, work, and shopping), the differential land use of cities is a key to comprehending neighborhood crime, and, by implication, disorder patterns. The effects of concentrated resource disadvantage and residential instability on disorder should thus be considered in concert with structural characteristics such as density, street activity, and commercial land use.

Structural constraints notwithstanding, one might view human agency as central to the explanation of disorder. In this view, it is not only the material circumstances or ecological structures that residents face, but the challenge to organize themselves to achieve shared public ends. We adopt the formulation of Janowitz (1975, pp. 82, 87) and refer to social control as the capacity of a social unit to regulate itself according to desired principles—to realize collective, as opposed to forced, goals. Hence, social control should not be equated with repression or forced conformity. Similar to Bursik (1988), our strategy also highlights variations in social control across ecological units rather than elevating solidarity or affective identity to the major definitional criteria of neighborhood (see also Tilly 1973, p.

212). When formulated in this way, the dimensions of social control are analytically separable not only from possible structural antecedents (e.g., poverty, instability) and effects (e.g., disorder, crime) but from the definition and operationalization of the units of analysis.

Building on Janowitz's (1975) conception of social control requires that we explicitly note the assumption of relative consensus; namely, one of the most central of common goals is the desire of community residents to live in safe environments free of predatory crime and disorder. There is no evidence of which we are aware showing public approval of crime or disorder by any population group (Kornhauser 1978; Hearn 1997); if anything, the evidence suggests that residents of low-income, African-American, and high-crime neighborhoods are the *most* insistent on better police protection and demands for reducing violence (Skogan and Hartnett 1998). Although existential weariness in the inner city may lead to a greater tolerance of certain forms of deviance, it is precisely the acceptance of common standards by residents and even gang leaders themselves that underlies efforts to establish social order and safety—however unconventional those efforts may be (Sampson and Jeglum-Bartusch 1998). Indeed, Pattillo's (1998) revealing ethnography of a black middle-class neighborhood on the south side of Chicago found that the incorporation of gang members and drug dealers into the networks of law-abiding kin and neighbors thwarted efforts to rid the neighborhood of its criminal element. Yet in an interesting twist, Pattillo found that the leader of a major black gang was a long-time resident who engaged in multiple acts of social control (e.g., threats, monitoring) to keep the neighborhood free of street crime and signs of disorder (e.g., graffiti, vandalism, prostitution). Pattillo writes that both sides—the residents and the gang leaders—"spurn disorder, actively combat graffiti, and show disdain for activities that may invite negative attention, such as loitering or public fighting" (1998, p. 755). Her findings also highlight the important distinction between "crime" and "criminals"—ironically, even active participants in criminal networks seek to achieve some semblance of order in their neighborhoods of residence. The phenomenon of informal efforts to socially control local crime and disorder has long been reported in white working-class neighborhoods dominated by the mob (e.g., Whyte 1943).[6]

[6] Heeding Whyte (1943), we acknowledge that disorder is in some respects a misleading term if not properly contextualized, for obviously many elements of disorder do not involve disorganization or a chaotic pattern  Prostitution and drug dealing, e g., may follow quite explicit rules of street organization (Pattillo 1998)  By disorder, then, we refer not to disorganization but observable physical and social cues that are commonly perceived to disturb the civil and unencumbered use of public space (see Skogan 1990; Hearn 1997). Moreover, we focus on the ecological *concentration* of *multiple* dimensions of public disorder. A particular exhibit of graffiti might be considered an artistic expression, but the recurrent defacement of public property accompa-

In contrast to externally or formally induced actions (e.g., a police crack-down, housing code enforcement), our agency-oriented perspective on neighborhoods emphasizes the role of informal mechanisms by which residents initiate or achieve social control. Examples of indigenous informal social control relevant to reducing disorder include the willingness of residents to intervene to prevent acts such as truancy, drinking, vandalism, and street-corner disturbances (e.g., harassment, loitering, fighting). Importantly, however, actions of informal control need not involve direct confrontation or exclude the police or other formal channels of recourse. Recent ethnographic research has identified the creative ways in which socially organized communities react to disorder, including the establishment of "phone trees" among residents for calling the police upon observation of disorder; the organization of a group presence in court sentencing hearings for offenders caught defacing local public properties; voluntary "graffiti patrols" that log new incidents of disorder that are then presented to the police; and agitating for voting referendums to delicense bars where drug sales and disorder loom large (Carr 1998). The razing of a vacant "drug house" by housing authorities, if prompted by local complaints, would also fit this pattern. Ultimately, then, our perspective recognizes the articulation among the private (family), parochial (neighborhood), and formal (public) orders (Hunter 1985; Bursik and Grasmick 1993) but stresses the agency of residents in establishing these connections.

In short, we theorize informal social control as a dynamic process differentially activated across neighborhoods. Drawing on Sampson, Raudenbush, and Earls (1997), we propose an analogy between individual efficacy and neighborhood efficacy: both refer to the capacity for achieving an intended effect. At the neighborhood level, the shared willingness of local residents to intervene for the common good depends, in addition, on conditions of cohesion and mutual trust among neighbors. One is unlikely to take action in a neighborhood context where the rules are unclear and people mistrust one another. Personal ties and friendship are not sufficient; the private world of strong kinship ties may actually interfere with public trust and the expectation of collective responsibility for getting things done (Whyte 1943; Jacobs 1961, p. 82; Carr 1998). Attempting to transcend the traditional focus in urban systemic theory on personal ties, we thus define "collective efficacy" as the linkage of cohesion and mutual trust with shared expectations for intervening in support of neighborhood

---

nied by other physical signs of deterioration (e.g., broken windows, abandoned cars, trash in the streets) signals to most observers a perceived threat to public order. Neighborhood disorder is thus in the end a useful concept, not the least because it conjures up powerful visual images that cut across all population subgroups.

social control (Sampson et al. 1997).[7] Just as individuals vary in their capacity for efficacious action, so too do neighborhoods. And just as individual self-efficacy is situated relative to a particular task rather than global, our notion of collective efficacy here is conceptualized as relative to the task of maintaining order in public spaces.

To be sure, structural constraints and process-oriented mechanisms such as informal social control are not mutually exclusive. A more plausible account would view structural constraints and human agency as interrelated, jointly and reciprocally shaping social action (Sewell 1992). Within the limitations of our data, we therefore expect simultaneous contributions of structural characteristics and collective efficacy to the explanation of observed disorder and crime. Theory elaborated elsewhere (Sampson et al. 1997, p. 919) also suggests that concentrated resource disadvantage and residential instability are major structural conditions that undermine collective efficacy, in turn fostering increased crime and, by implication, public disorder. A theory of collective efficacy thus does not render structural constraints irrelevant; rather, it proposes mediating mechanisms while at the same time insisting on an independent role for agency in all corners of the social structure. We provide a strong test of the theory of collective efficacy by considering as well the simultaneous or reciprocal effect of crime itself on residents' sense of mutual trust and shared expectations for social control (Skogan 1990).

Research Strategy

Sharing an affinity with routine activity theory (Cohen and Felson 1979; Cohen, Kluegel, and Land 1981; Felson 1987; Brantingham and Brantingham 1995), the logic of our analytic approach diverges from a concern with the production of offenders as in the classic social-disorganization tradition of Shaw and McKay (1942; see also Bursik 1988). In the modern urban system, residents traverse the boundaries of multiple neighborhoods during the course of a day (Felson 1987), a problematic scenario for neighborhood theories seeking to explain contextual effects on individual differences in offending. By contrast, we are interested in how neighborhoods fare as units of control or guardianship over their own public spaces (Cohen et al. 1981)—regardless of where offenders may reside. The unit

[7] Note the affinity with Jacobs's (1961, p. 119) focus on the "self-government functions of city streets: to weave webs of public surveillance and thus to protect strangers as well as themselves; to grow networks of small-scale, everyday public life and thus of trust and social control; and to help assimilate children into reasonably responsible and tolerant city life." Social order in this vision does not require personal friendship or kinship ties but rather collective expectations for action in the public sphere

of analysis is thus the neighborhood, and our phenomenon of interest is the physical and social disorder within its purview.[8]

Relatedly, the theoretical logic of collective efficacy focuses foremost on activity patterns that can be visibly observed. Here is where the theory links most naturally to the SSO method. Disputes among acquaintances and domestic violence that occurs indoors, for example, are by their very nature less amenable to public surveillance and sanctioning. They very likely are also perceived (albeit incorrectly) as less threatening to the common good. Collective efficacy is thus particularly relevant to explaining the incidence of crime and disorder in public spaces, and to crimes like robbery and burglary that typically elicit target selection decisions based on visual cues.[9] It follows that if the theory of collective efficacy is valid, it should be able to explain variations in independently collected measures of disorder in public spaces. We therefore introduce the method of systematic observation as a critical test case for the hypothesis that collective efficacy inhibits neighborhood disorder.

Finally, by implication, our theory offers a different way to think about the question of ecological "comorbidity"—the association between public disorder and predatory crime, especially violence. The "broken windows" literature sees disorder as a fundamental cause of crime (Skogan 1990, p. 75; Kelling and Coles 1996); if true, the hypothesized association of structural characteristics and collective efficacy with crime and violence ought to be largely mediated by social disorder. The alternative hypothesis we offer is that disorder is a manifestation of crime-relevant mechanisms and that collective efficacy should reduce disorder *and* violence by disempowering the forces that produce both. Our theory suggests also that structural constraints such as resource disadvantage and mixed land use account for both crime and disorder simultaneously. We thus test whether neighborhood disorder is an essential link in the ecological pathway that leads to predatory crime or rather a spuriously related construct rooted in structural constraints and collective social processes.

---

[8] Parenthetically, we believe that the routine activity insight has not sufficiently penetrated "neighborhood effects" research in sociology. A thought experiment reveals the possibility that 100% of the incidents of crime (or disorder) could be concentrated in one neighborhood (e.g., stimulated by opportunities and low surveillance) but with constancy in offending across neighborhoods (i.e., each contributing an equal rate of offenders). The former would suggest a neighborhood effect and the latter none. More generally, neighborhood research on behaviors such as academic achievement, employment, and drug use is problematic to the extent that the behaviors of interest are embedded in multiple ecological contexts (e.g., schools, parks, businesses) outside the respondent's neighborhood of residence. Clearly, a plausible theory of the ecology of the phenomenon under study is the first order of business.

[9] See Wright and Decker (1997). Note also that a surprising proportion of property crimes (almost 50%) occurs in public places (Felson 1987)

TABLE 1

DISTRIBUTION OF SAMPLED CHICAGO CENSUS TRACTS AND NEIGHBORHOOD CLUSTERS

| | SES | | |
|---|---|---|---|
| RACE/ETHNICITY | Low | Medium | High |
| 75% black or more ....... .... .... ......... .. . ....... ... ...... | 31 (9) | 10 (4) | 9 (4) |
| 75% white or more .... . ............... .. . ............ . ....... | 0 (0) | 7 (4) | 18 (8) |
| 75% latino or more .. ............... .. . ............... ........... | 12 (4) | 12 (4) | 0 (0) |
| 20% latino or more/20% white or more ... .. .......... . | 11 (4) | 14 (5) | 10 (4) |
| 20% latino or more/20% black or more .......... .. .. | 7 (4) | 7 (4) | 0 (0) |
| 20% black or more/20% white or more ......... . ..... | 3 (2) | 4 (4) | 10 (4) |
| Other heterogeneous ....... .. ............. . ............ .. ...... | 8 (4) | 14 (4) | 9 (4) |
| Total ... ............. . .............. .. . ............. . ... ..... | 72 (27) | 68 (29) | 56 (24) |

NOTE.—SES was defined by a six-item index that summed standardized census-based measures of median income, % college educated, % with household income over $50,000, % families below the poverty line, % on public assistance, and % with household income less than $5,000 The last three items were reverse coded. Neighborhood clusters are given in parentheses.

RESEARCH DESIGN

Between June and October 1995, observers trained at the National Opinion Research Center (NORC) drove a sport utility vehicle (SUV) at a rate of five miles per hour down every street in 196 Chicago census tracts. These tracts were selected from a stratified probability sample to maximize variation by race/ethnicity and SES. As part of a larger study, 343 "neighborhood clusters" (NCs) representing combinations of all 865 census tracts in Chicago were first stratified by seven categories of race/ethnic mix and three levels of SES.[10] Within strata, 80 NCs were then sampled for intensive study with the aim of obtaining a near balanced design, eliminating the confounding between ethnic mix and SES (table 1). However, in Chicago as in many other cities, ecological sorting by race and class results in a sample with some empty cells—low SES, predominantly European-American; high SES, predominantly Latino; and high SES, mixed Latino and black. Also, the largest stratum was low SES and predominantly African-American and contained 77 NCs, generally characterized by concentrated poverty, racial segregation, and other forms of disadvantage. The final design randomly sampled four NCs within cells that had

[10] Containing about 8,000 people, NCs are composed of geographically contiguous census tracts with similar distributions on key census indicators (e.g., race, SES, family structure, housing). We also used geographic boundaries (e.g., railroad tracks, parks, and freeways) and our knowledge of Chicago's neighborhoods as guides for constructing NCs. For more details, see Sampson et al. (1997, p. 924).

at least four, all NCs within cells having fewer than four, with an oversampling of the largest and most disadvantaged cell. Although reflecting the reality of segregation by race and class, the sampled NCs nonetheless tap the maximal existing variation across these dimensions of stratification. The distribution of the 196 census tracts embedded within the set of 80 sampled neighborhood clusters is shown in table 1.

The geographic unit of recorded observation within the sampled NCs and tracts was the face block: the block segment on one side of a street. For example, the buildings across the street from one another on any city block comprised two separate units of observation. At each intersection, a unique geographic identification code was assigned so that adjacent block faces could be pieced together to form higher levels of aggregation desired by theory or as suggested by patterns in the data. To observe each block face, the NORC team fielded a driver, a videographer, and two observers. As the SUV was driven down the street, a pair of video recorders, one located on each side of the SUV, captured social activities and physical features of both face blocks simultaneously. At the same time, two trained observers, one on each side of the SUV, recorded their observations onto an observer log for each block face. The observers added commentary when relevant (e.g., about unusual events such as an accident or drug bust) by speaking into the videotape audio. Face blocks were observed and videotaped between the hours of 7 A.M. and 7 P.M.[11] Applying these procedures, the SSO team produced videotapes, observer logs, and audiotapes for every face block in each of the 80 sampled NCs.

In all, 23,816 face blocks were observed and videotaped, for an average of 298 per NC and 120 per tract.[12] The data collected from the 23,816 observer logs focus mainly on land use, traffic, the physical condition of buildings, and evidence of physical disorder. Unlike the observer logs, which could be directly entered into machine-readable data files, the videotapes required the expensive and time-consuming task of first viewing and then coding. We thus selected a random subsample of all face blocks for coding. In those NCs consisting of 150 or fewer face blocks, all face blocks were coded; in the remaining NCs, sample sizes were selected to

[11] Although it would be desirable to assess disorder during the nighttime hours, a pretest confirmed that this was not feasible with current videotaping technology (or with the naked eye).

[12] The coding protocols for the observation logs and videotapes were first piloted in five census tracts from another city (Boston). After some modification, the protocols were then evaluated in a series of pretests in nonsampled neighborhoods of Chicago. Based on this experience, all major operational decisions were finalized. Coders were then trained in multiple sessions, including an intercoder reliability training where 90 face blocks were independently double coded, differences resolved, and coding procedures revised.

approximate a balanced design as closely as possible in order to maximize statistical power for comparisons of NCs and tracts. A total of 15,141 face blocks were sampled for videotape coding, an average of 189 per NC and 77 per tract. From the videotapes, 126 variables were coded, including detailed information on physical conditions, housing characteristics, businesses, and social interactions occurring on each face block (NORC 1995). As a check on quality control, new observers recoded a random 10% of all coded face blocks, and the results compared. This test produced over 98% agreement (NORC 1995; Carter, Dougherty, and Grigorian 1996).

Measures

Although some of the items were measured initially on an ordinal scale, the data behaved essentially as dichotomous items and so were coded for analysis as 1 = presence and 0 = absence of the indicator of disorder (Raudenbush and Sampson 1999b). The first scale is based on 10 items intended to capture the presence or absence of *physical* disorder. In declining order of observed frequency, the scale items include the presence or absence of *cigarettes or cigars* in the street or gutter (no = 6,815; yes = 16,758); *garbage or litter* on street or sidewalk (no = 11,680; yes = 11,925); *empty beer bottles* visible in the street (no = 17,653; yes = 5,870); *tagging graffiti* (no = 12,859; yes = 2,252); *graffiti painted over* (no = 13,390; yes = 1,721); *gang graffiti* (no = 14,138; yes = 973); *abandoned cars* (no = 22,782; yes = 806); *condoms* on sidewalk (no = 23,331; yes = 231); *needles/syringes* on sidewalk (no = 23,392; yes = 173); and *political message graffiti* (no = 15,097; yes = 14). The variation in sample size reflects the fact that six of the ten items were taken from the observer logs and thus have nearly complete data. The other four variables were derived from the videotapes and thus are based on the reduced subsample selected for coding. As expected, less serious indicators of disorder in public spaces (e.g., presence of cigarettes and garbage) arise more frequently than do indicators that might be regarded as more serious (e.g., drug paraphernalia and gang graffiti), with the presence of beer bottles arising with moderate frequency. Political graffiti was very rarely observed, even though it is not necessarily an indicator of severe disorder.[13]

[13] Detailed memos on the recognition of graffiti and gang insignia guided the training of coders (NORC 1995) Graffiti was classified by type—tagging, gang, and political—based on investigator guidelines, gang research in Chicago (Spergel 1995), and internal Chicago police memos on gang identification. Tag graffiti is identified by stylized forms such as block letter art or by attempts to create some form of visual expression. Gang graffiti is ordinarily distinguished by the absence of tag art and usually includes messages that refer to the name of a rival gang. Gang symbols in Chicago are long-standing and widely recognized and typically involve a combination of stars, crowns, emblems, and specific colors that distinguish among gangs (Spergel 1995; NORC

Direct evidence of *social* disorder was coded from the videotapes. The scale items include presence or absence of *adults loitering or congregating*[14] (no = 14,250; yes = 861); *drinking alcohol in public* (no = 15,075; yes = 36); *peer group with gang indicators* present (no = 15,091; yes = 20); *public intoxication* (no = 15,093; yes = 18); *adults fighting* or arguing in a hostile manner (no = 15,099; yes = 12); *selling drugs* (no = 15,099; yes = 12); and *prostitutes* on the street (no = 15,100; yes = 11). Most items in the disorder scale bear a conceptual affinity with concurrent "crime" in the sense of violation. Moreover, the frequency distribution tells us that, like crime, social disorder is quite rare—at least during the daylight hours. If social disorder resembles crime, indicators of the former should also be present far less frequently than indicators of physical disorder. The frequencies support this inference too. Activities that might be viewed as indicating more serious disorder (prostitution, drug selling, adults fighting) are especially rare. Indicators that are somewhat less serious are also relatively less rare (drinking alcohol, presence of peer gangs), though they remain rare overall relative to physical disorder. One item—adults loitering—occurred with much higher frequency than did any other item.

The frequency distribution of items suggests that the physical disorder scale will behave better as an ecological measure than will the social disorder scale. It not only has more items (10 vs. 7) but more important, the physical disorder items range widely; several occur with large frequency, several others with modest frequency, and several are comparatively rare. By contrast, the social disorder indicators occur with rare frequency except for adults loitering or congregating. The social disorder scale may thus be dominated by a single item of relatively low frequency, leading to an overall lack of between-neighborhood reliability.

Raudenbush and Sampson (1999*b*) adapted tools found useful in psychometrics to the problem of evaluating the systematic social observation of disorder. There are multiple components of measurement error that they addressed: (*a*) item inconsistency within a face block, (*b*) face-block variation within larger geographic units, and (*c*) temporal variation. The latter is a particular problem in measuring social disorder (Perkins et al. 1992; Taylor 1999). The probability of finding adults loitering or drinking, of finding peer gangs hanging out, or of seeing prostitution or drug deals

will depend on the time of day in which a face block is observed. Thus it is necessary to estimate and adjust for time of day. Fortunately, because face blocks were assessed between the hours of 7 A.M. and 7 P.M., there is considerable temporal variation in the time of observation within each census tract. We also need to allow for randomly missing data, because only a random sample of face blocks yielded data coded from videotapes.

To achieve these goals, we modify the three-level hierarchical regression model described by Raudenbush and Sampson (1999b).[15] The level-1 units are scale-item responses within face blocks, the level-2 units are face blocks, and the level-3 units are the 196 census tracts embedded within the sampled 80 NCs. In this article, we operationalize neighborhood using the 196 census tracts rather than NCs for three important reasons. First, previous research has argued for the smallest level of aggregation possible in measuring observed disorder, owing to considerable variability block-to-block within larger ecological units (Perkins et al. 1992; Taylor 1997). Second, by dropping down to the tract level from NCs, we more than double the neighborhood-level degrees of freedom, thus providing more statistical power to detect between-area variations. The number of face blocks observed within each tract was large enough to produce reliable measures of mean tract-level differences; the same was *not* true for block groups or block-level measures. Third, census tracts provide the additional information necessary to address the well-known multicollinearity among ecological variables (Land et al. 1990). At the tract level, the overlap among variables is considerably less than for NCs.

The appendix presents the three-level statistical model and measurement properties of the physical disorder and social disorder scales that are the main substantive focus of the ensuing analysis. The results show that the two disorder scales are highly reliable at the ecological level. Differences between neighborhoods in their aggregated disorder scores can be interpreted as expected differences in the log-odds of finding disorder on a typical item in the scale, adjusted for time-of-day effects in observation. The scales used below are thus meaningfully interpretable and arguably a linear (interval) scale appropriate for analysis via standard linear models.

INDEPENDENT DATA COLLECTION

To assess the theoretical framework on collective efficacy, structural constraints, and observed public disorder, we examine independent sources of data collected from a survey, police records, vital statistics, and the

---

[15] The initial methodological work for the SSO model was conducted on the 80 NCs, generally with similar measurement results (for details, see Raudenbush and Sampson 1999b).

census. First, over 4,000 households within the 196 tracts and 80 NCs were selected according to a multistage probability sample (see Sampson et al. 1997 for details). Within each household, a randomly chosen adult was interviewed in late 1994 and in the first eight months of 1995 concerning conditions and social relationships in the local neighborhood. The final sample size was 3,864, reflecting a response rate of 78%.[16]

Derived from the theoretical strategy outlined earlier, we examine four constructs from the survey—disorder, cohesion, control, and crime—aggregated to the tract level. Predatory crime was measured from respondents' reports of whether they (or any member of the household) had experienced within the past 6 months (a) a violent victimization in the neighborhood or (b) a household burglary or theft victimization. Approximately 5% of respondents reported a violent crime and 16% a household crime. The multi-item disorder scale taps how much a problem ("a big problem," "somewhat of a problem," "not a problem") residents rated the presence in the neighborhood of social incivilities (drinking in public, selling or using drugs, teenagers causing a disturbance) and physical incivilities (litter, graffiti, vacant housing). The reliability of the combined scale of perceived disorder at the tract level is .83 (for details on aggregate-level reliability see Raudenbush and Sampson 1999a).

A measure of shared expectations for informal social control was represented by a five-item Likert-type scale. Residents were asked about the likelihood ("Would you say it is very likely, likely, neither likely nor unlikely, unlikely, or very unlikely?") that their neighbors could be counted on to take action ("do something") if (a) children were skipping school and hanging out on a street corner, (b) children were spray painting graffiti on a local building, (c) children were showing disrespect to an adult, (d) a fight broke out in front of their house, and (e) the fire station closest to home was threatened with budget cuts. "Social cohesion/trust" was represented by five conceptually related items. Respondents were asked how strongly they agreed (on a 5-point scale) that "People around here are willing to help their neighbors," "This is a close-knit neighborhood," "People in this neighborhood can be trusted," and, reverse coded, "People in this neighborhood generally don't get along with each other" and "People in this neighborhood do not share the same values." Social cohesion and informal social control were correlated at r = .68 ($P < .01$), suggesting that the collective willingness to intervene in the neighborhood is enhanced under conditions of mutual trust and cohesion. Following Sampson et al. (1997), the two measures were combined to create a more parsi-

[16] The average number of survey respondents per tract was 20. One primarily nonresidential tract was not available for further analysis because it had only one survey respondent (with missing data).

monious and readily interpretable measure of "collective efficacy" with strong theoretical connections to disorder reduction. The aggregate reliability is .68 and .80 at the tract and NC levels, respectively.

The second set of data taps ecological variations in crime independent of the survey. The main source consists of incidents (not arrests) of homicide, robbery, and burglary in the years 1993 and 1995, geocoded from records of the Chicago Police Department and aggregated to the census tract of occurrence.[17] To provide some purchase on controlling for prior sources of crime not captured in our measured variables, 1993 data are employed. The crimes of homicide, robbery, and burglary are generally well reported and provide insight on variation in predatory crimes involving both property and persons. Comparatively, however, we place the most confidence in homicide as it is the most reliably measured of all crimes and does not suffer major reporting limitations. Because homicide is so rare—50% of the tracts had no incidents—we analyze the raw count of homicide using a negative binomial regression.[18] By contrast, all neighborhoods had recorded incidents of burglary and robbery, with a mean number around 40. We analyze the log of the robbery rate per 100,000 persons (mean = 6.90) and the log of the burglary rate per 100,000 households (mean = 8.32) as outcomes, both of which approximate very well a normal distribution. Overall, this research strategy provides five indicators of crime-rate variation covering property and violent crimes, and as measured in both surveys and police incident data. In addition, a person-based measure of homicide victimization derived from vital statistics rather than police records is employed in later analysis (described below).

The third source of independent data was culled from 1990 census data at the tract level. Three indexes of neighborhood structural differentiation are examined based on prior theory (Wilson 1987; Sampson et al. 1997) and research analyzing census data in Chicago over three decades (Morenoff and Sampson 1997). To reduce the dimensionality of the data, an alpha-scoring factor analysis with an oblique factor rotation was performed on the tract level, replicating the results of Sampson et al. (1997, p. 920) at the NC level. *Concentrated disadvantage* represents an economic disadvantage factor in racially segregated urban neighborhoods that was dominated by high loadings (> .8) for poverty, public assistance, unemployment, and female-headed families. Percentage of black residents was also linked to this dimension, although to a lesser extent (.62). Hence, this factor reflects the neighborhood concentration of resource disadvantage,

---

[17] We owe a debt of gratitude to Richard Block for providing the raw data.

[18] Poisson regression yielded evidence of overdispersion and thus the desirability of a negative binomial model, which incorporates a variance parameter to represent heterogeneity across tracts.

to which African-Americans and single-parent families with children are disproportionately exposed (Wilson 1987; Land et al. 1990). The second factor captures areas of *concentrated immigration*. The variables that loaded high on this dimension were percentage Latino (.95), percentage foreign-born (.73), and, to a lesser extent, density of children (percentage of persons ages 6–15), which loaded at .6. The third factor was dominated by two variables with very high (> .8) loadings—percentage living in the same house as five years earlier, and percentage of owner-occupied homes. The emergence of a *residential stability* factor is consistent with much past research (Kasarda and Janowitz 1974). Using factor loadings as weights, summary scales were created to reflect the three dimensions.[19]

As noted earlier, the routine activities approach (Cohen and Felson 1979; Cohen et al. 1981; Wikström 1991) suggests that ecological characteristics of neighborhoods reflect opportunities for crime and bear on the ability of residents to engage in guardianship. We therefore control for two ecological constructs emphasized in this approach—land use and density. Using census data, we control for the number of persons per square kilometer in the tract (mean = 7,530). Neighborhoods with more people per unit of space may generate greater anonymity and persons in public, making it harder for residents to maintain informal social control over public space.[20] The second control for neighborhood ecology is land use, defined as the proportion of face blocks in the tract that contain mixed residential and commercial activity (mean = 25%). Mixed land use has been shown to be a robust but understudied correlate of crime and disorder (see, e.g., Wikström 1991; Taylor 1995) and is theoretically relevant to understanding collective efficacy as well. It may be, for example, that the capacity of residents to achieve common purpose is limited not because of lack of internal effort but simply the structural constraint imposed by the density of commercial traffic and land-use patterns inhospitable to social interaction and surveillance. The mixed land-use variable is well measured, constructed from the complete set of 23,816 observational coding logs.

## SOURCES AND CONSEQUENCES OF DISORDER

Observed disorder is correlated at the tract level with those constructs measured in the community survey and other independent sources (census

---

[19] A principal components analysis produced equivalent results (e.g., the tract-level correlation between a principal components and oblique rotation for the poverty scale was .99).

[20] We examined several other indicators of density and street activity, including housing density and the number of workers taking public transportation per capita. However, these were either highly correlated with population density or added little to the explanatory models.

data, official police records) most theoretically linked to it. For example, a healthy Pearson correlation of r = .56 (*P* < .01) emerges between SSO social disorder and social disorder measured in the community survey. SSO physical disorder is correlated .55 (*P* < .01) with survey-reported physical disorder. It is noteworthy that the correlation between residents' perceptions of physical disorder and SSO physical disorder is nearly identical to the correlation between survey-reported social disorder and SSO social disorder, given that physical disorder is relatively stable over time and social disorder reflects events that occur randomly in time. SSO disorder is also moderately strongly correlated with the survey measures of collective efficacy (r = −.49 for social disorder; r = −.47 for physical disorder; *P* < .01) in the direction expected.

Turning next to correlations with sociodemographic composition, SSO physical disorder is significantly related to census measures of concentrated poverty (r = .50) and immigrant concentration (r = .39). The relationship with residential stability is weak and insignificant. Furthermore, physical disorder measured in the SSO is only moderately correlated with rates of predatory crime measured by police-recorded rates of homicide (r = .27), robbery (r = .45), and burglary (r = .24). The relationships with survey-reported victimization are weaker: r = .21 for violence, and r = .06, NS, for burglary. A similar pattern of correlation appears with respect to social disorder. The bivariate results thus suggest that SSO measures of disorder have reasonably consistent relationships with theoretically linked explanatory factors derived from the neighborhood survey and census. Notably, however, the correlations of SSO disorder with crime rates, although positive, are not at the levels one might expect from the broken windows thesis.

Table 2 begins to unpack the sources and consequences of observed disorder across our sample of 195 Chicago census tracts in theoretically specified multivariate models. We examine first the weighted least squares (WLS)[21] regression of physical and social disorder measured by SSO on dimensions of structural differentiation, collective efficacy, density, and land use.[22] With approximately 50% of the variance explained, note the

[21] As noted above, the disorder scales form a linear and near-normally distributed metric. However, because the number of blocks used to create the scales varied by tract, WLS regression was used to induce homoscedasticity of error variances. Each case was weighted by the square root of the unweighted number of assessed blocks, giving more weight to tracts with a larger sample of coded data (see Hanushek and Jackson 1977, pp. 143, 152).

[22] Using the three-level hierarchical logistic regression model for uncertainty in measuring observed disorder elucidated in the appendix. the tract-level measures account for the multiple sources of measurement error identified (e.g., temporal, missing data). In this framework, empirical Bayes (EB) residuals are analyzed, defined as the least-squares residuals regressed toward zero by a factor proportional to their unreliability

TABLE 2

WLS TRACT-LEVEL REGRESSION OF SSO PHYSICAL AND SOCIAL DISORDER ON CENSUS
STRUCTURAL CHARACTERISTICS, SSO LAND USE, AND SURVEY-BASED COLLECTIVE
EFFICACY

| | SYSTEMATIC SOCIAL OBSERVATION | | | |
| | Physical Disorder | | Social Disorder | |
| | B | t-ratio | B | t-ratio |
|---|---|---|---|---|
| Concentrated disadvantage . ... ...... . | .44 | 7.56** | 40 | 6 58** |
| Residential stability ..................... ........... | .01 | 12 | − 02 | −.36 |
| Immigrant concentration ......... ........... | .28 | 4.99** | .19 | 3.20** |
| Population density ...... .. .. ........ .. . . | 03 | .57 | .01 | .11 |
| Mixed land use . ... .. ............ ... . . | .18 | 3.19** | .24 | 4.08** |
| Collective efficacy .. .. .. ............ ... .. . | −.19 | −2.54* | −.21 | −2.76** |
| Adjusted $R^2$ . .. ..................................... | .52 | | .48 | |

NOTE.—$N$ = 195
* $P$ < 05
** $P$ < 01.

strong association of disorder with concentrated disadvantage and immi-
grant concentration. The standardized coefficients for concentrated disad-
vantage are by far the largest in predicting physical disorder (B = .44; t-
ratio = 7.56) and social disorder (B = .40; t-ratio = 6.58). After accounting
for structural aspects of neighborhood differentiation by class and race/
ethnicity, we also observe significant relationships with land use. Neigh-
borhoods with mixed residential and commercial development exhibit
higher levels of both physical and social disorder, regardless of sociodemo-
graphic characteristics. Furthermore, the estimated association of collec-
tive efficacy with both forms of disorder is consistently negative and sig-
nificant.

The data thus far suggest that structural constraints matter greatly in
predicting disorder. Moreover, collective efficacy with respect to the social
control of public space appears to inhibit the incidence of observed disor-
der, whether physical or social in nature. That this relationship is some-
what stronger with social disorder (unstandardized b = −.81) than physi-
cal disorder (b = −.65) is consistent with the theoretical underpinnings

---

(Bryk and Raudenbush 1992, chap. 3). Similarly, the collective efficacy and perceived
disorder scores for each tract were computed using the EB residuals, also accounting
for measurement error and missing data (see also Raudenbush and Sampson 1999a).
Using EB residuals as explanatory variables corrects for bias in regression coefficients
resulting from measurement error (Whittemore 1989). Although more precise, the EB
results nonetheless mirror the results using simple scale averages.

TABLE 3

WLS Tract-Level Regression of SSO Physical and Social Disorder on Prior
Homicide Rate, Perceived Disorder, and Collective Efficacy

| | SYSTEMATIC SOCIAL OBSERVATION | | | |
| --- | --- | --- | --- | --- |
| | Physical Disorder | | Social Disorder | |
| | B | t-ratio | B | t-ratio |
| Survey physical disorder ...... . ......... . | .41 | 5.89** | ... | |
| Survey social disorder .......... ......... ... | .. | .. | 41 | 5.36** |
| Prior (1993) homicide ....... .. ...... .. ...... | .15 | 2.59* | .08 | 1.39 |
| Collective efficacy . .......... . ...... .......... . | −.26 | −3.69** | −.25 | −3.28** |
| Adjusted $R^2$ ..... ........... ......... ......... . | 42 | | .39 | |

Note —$N$ = 195
* $P < 05$.
** $P < 01$.

of collective efficacy. Clearly, however, concentrated disadvantage is the
single most important predictor of disorder in Chicago neighborhoods.
The structural origins of disorder cannot be ignored (Hunter 1985).[23]

Table 3 presents a methodologically oriented test of the association of
collective efficacy with observed disorder. It may be that respondents'
perceptions of disorder in the neighborhood colored their judgments about
cohesion and control. That is, informants might reasonably infer low col-
lective efficacy from their perceptions of disorder in the neighborhood.
Because survey disorder (combined) is correlated −.63 with collective ef-
ficacy and .62 with SSO disorder ($P < .01$), controlling its effects serves
as a strict test. After all, in order for disorder to confound the efficacy
scale, it must be mediated by the perceptions of the reporter. Accordingly,
we estimate the net effect of efficacy on disorder *after* resident perceptions
of disorder have been removed. Because the vast majority of the survey
was completed before the SSO, this control might also be interpreted as
estimating the effects of collective efficacy on changes in disorder. To pro-
vide another guard against the endogeneity of collective efficacy, we con-
trol for prior violence as measured by the 1993 homicide rate. Total vio-

[23] All regression analyses were examined for influential observations using a variety
of diagnostics (e.g., Cook's D; leverage scores) No one tract exerted a disproportionate
influence on the coefficient estimates after appropriate (e.g., log) transformations. We
also examined various diagnostics for multicollinearity. At the tract level, this did not
turn out to be a problem. In fact, all correlations among predictor variables were less
than .50, and the variance inflation factors (VIF) were all less than 2.5, well below
those levels traditionally thought to be of concern (see Belsley, Kuh, and Welsch 1980).

lence showed similar but weaker results. Thus, both lagged homicide and concurrent perceived disorder are used to adjust the levels of observed disorder, disaggregated in both the survey and SSO measures by type.

The data in table 3 are very clear in showing a consistent negative relationship of collective efficacy with SSO disorder, both physical and social. The estimated net effect of collective efficacy on physical disorder is B = −.26 (t-ratio = −3.69), larger than the (significant) direct effect of prior homicide (B = .15). The effect of collective efficacy on social disorder is the same magnitude (B = −.25; $P < .01$). Not surprisingly, the direct association of perceived with observed disorder is larger, regardless of whether physical or social in nature. The key result, however, is that whether we control for sociodemographic characteristics from the census, prior violence, or even perceived disorder, the data show a persistent negative association of collective efficacy with the independently observed incidence of disorder.[24] Because the results in tables 2–3 are so similar for physical and social disorder, and because the two scales are highly correlated (r = .71 for EB residuals), we combine them into a summary index of SSO disorder for the remaining analysis.

### "Broken Windows" Revisited

In the next set of tests, we turn to the consequences of disorder. A fundamental thesis of "broken windows" is that observed disorder directly causes predatory or "serious" crime. An alternative interpretation is that disorder and crime are both the products of weakened social controls and structural antecedents. Table 4 provides evidence that adjudicates between these two scenarios using survey-reported victimization aggregated to the neighborhood level. Model 1 in the top half of the table, with no controls for structural characteristics or collective efficacy, shows that SSO disorder *is* positively linked to survey-reported violence (t-ratio = 2.50; $P < .05$), although only 5% of the overall variance is explained. Interestingly, officially measured rates of prior violence are not significantly related to later survey-reported violence. By contrast, 1993 burglary rates predict later survey-reported burglary, while SSO disorder is seen to be unimportant. This reveals that disorder, systematically measured with observation, is positively but rather weakly related to survey-reported neighborhood victimization.

---

[24] When the census structural factors are added to table 3, the significant link between collective efficacy and SSO disorder is maintained. By contrast, the survey-based measure of disorder, which is correlated at r = .63 ($P < .01$) with concentrated disadvantage, is rendered insignificant. Unlike collective efficacy, then, perceptions of disorder appear to be confounded with population composition.

TABLE 4

WLS TRACT-LEVEL REGRESSION OF SURVEY-REPORTED VICTIMIZATION IN THE
NEIGHBORHOOD

| | SURVEY-REPORTED VICTIMIZATION, 1995 | | | |
| --- | --- | --- | --- | --- |
| | Personal Violence | | Household Burglary | |
| | B | t-ratio | B | t-ratio |
| Model 1: | | | | |
| SSO disorder ......... ......... . . ...... . | 20 | 2 50* | .08 | 1.10 |
| Prior (1993) crime ........................ . . | 08 | 1 03 | .22 | 3.16** |
| Adjusted $R^2$ ..... ......... .......... .. . ....... | | 05 | | .05 |
| Model 2: | | | | |
| SSO disorder ... .. ........... ..... .. .. .. | .12 | 1.11 | −.02 | −.16 |
| Prior (1993) crime ....._......... .. .. ..... | .00 | 02 | .24 | 3.15** |
| Concentrated disadvantage .. . ........ | −.03 | −.25 | −.14 | −1.51 |
| Residential stability . ........... .. ......... . | −.04 | −.44 | .07 | .74 |
| Immigrant concentration . ........... .. | 10 | 1 16 | .23 | 2 75** |
| Population density ......._ .. ......... .. | −.16 | −1.65 | −.14 | −1.53 |
| Mixed land use .......... _ ...... . ... .... | −.17 | −2 05* | .04 | .45 |
| Collective efficacy .... ........... .. .. ....... | −.37 | −3.59** | −.22 | −2.16* |
| Adjusted $R^2$ ......... ........._ .. ......... | | .13 | | .13 |

NOTE —$N$ = 195 For 1995 survey violence, the control for prior crime is the logged incident rate in 1993 of murder, rape, robbery, and aggravated assault per capita For burglary victimization, the control is the logged rate of 1993 burglaries per household
* $P < .05$
** $P < .01$.

In the bottom half of table 4 (model 2), we reestimate the effects of disorder and prior crime after adding controls for structural antecedents and collective efficacy. The results are consistent and point to a spurious association of disorder with predatory crime. In the case of violent victimization, the coefficient for SSO disorder is cut appreciably and reduced to insignificance. Collective efficacy, on the other hand, is by far the largest predictor of violent victimization (B = −.37; $t$-ratio = −3.59), absorbing the prior effects of concentrated poverty, residential stability, and immigrant concentration (see also Sampson et al. 1997). In fact, controlling for collective efficacy alone eliminated the effect of disorder. It is noteworthy as well that collective efficacy is associated with lower rates of violence regardless of sociodemographic composition and crime-linked mechanisms—namely, prior violence and observed neighborhood disorder. Columns 3 and 4 reveal a similar picture for household burglary victimization. Here we see that collective efficacy is linked to significantly lower burglary rates (B = −.22). Latino neighborhoods undergoing immigration flows experience higher rates of survey-reported household victimization.

SSO disorder showed almost no explanatory power in the first place ($t$-ratio = 1.10), and its coefficient in the full model is reduced to near zero.

We conducted several other tests to assess the robustness of the results for survey-reported victimization. Because violent victimization is rare, the measure was highly skewed—some 50% of the areas yielded no incidents. We therefore re-estimated the violent victimization model using logistic regression, with neighborhoods experiencing no incidents coded "0" and neighborhoods with one or more coded "1". The estimated effect of collective efficacy was still the largest ($t$-ratio = −3.22), and SSO disorder remained insignificant ($t$-ratio = .83). We also examined a summary measure of total reported victimization in the neighborhood, and in addition, we addressed shared method variance by controlling for perceived disorder. Reported victimization is correlated significantly with perceived disorder and may thereby confound informant reports of neighborhood collective efficacy. Under this specification, the effect of collective efficacy remained significant ($t$-ratio = −3.79), while the SSO-disorder link remained insignificant. Collective efficacy also predicted survey-reported violence (B = −.39) controlling for perceived disorder, whereas perceived disorder exhibited a null effect. Finally, we examined SSO physical disorder and social disorder in separate models to see if there were differential effects; neither measure emerged as a significant predictor. To this point then, it cannot reasonably be concluded that disorder is a proximate or strong mechanism explaining crime-rate variation. Rather, the results suggest a direct association of collective efficacy with lower rates of predatory victimization, unmediated by social disorder.

Table 5 turns to the other set of independently measured outcomes— officially recorded incidents of homicide, robbery, and burglary in 1995. At the bivariate level (model 1), SSO disorder predicts all three types of official crime, especially robbery rates (B = .48; $t$-ratio = 7.67). Structural antecedents, land use, and collective efficacy are introduced next as covariates at the multivariate level (model 2). For burglary, we see that SSO disorder drops out of the picture. In predicting homicide, the estimated coefficient for disorder also drops (by half) but retains significance. Only for robbery does the estimated effect of disorder remain substantively large, with a standardized coefficient of .31 ($t$-ratio = 3.86). Across all three crimes, the consistent predictors are concentrated poverty and collective efficacy. The latter's coefficient for homicide is −1.97 ($t$-ratio = −4.12); further calculation indicates that a standard deviation increase in collective efficacy was associated with a 35% decrease in the expected homicide rate after adjusting for prior homicide, disorder, and structural antecedents. A similar outcome obtains for collective efficacy in predicting rates of robbery and burglary. In fact, collective efficacy has the largest standardized effect estimate on robbery (B = −.35) and the second largest

## TABLE 5

### TRACT-LEVEL REGRESSION OF POLICE-RECORDED CRIME INCIDENTS IN 1995

| | POLICE-RECORDED INCIDENTS, 1995 | | | | | |
| | Homicide Counts | | Log Robbery Rate | | Log Burglary Rate | |
| | Coefficient | z-ratio | B | t-ratio | B | t-ratio |
|---|---|---|---|---|---|---|
| **Model 1:** | | | | | | |
| SSO disorder ... .......... .. ........ ... | 27 | 5 23** | .48 | 7.67** | .26 | 3.79** |
| **Model 2:** | | | | | | |
| SSO disorder ...... ....... .. ...... . | 13 | 2.17* | .31 | 3.86** | .07 | .75 |
| Concentrated disadvantage .. .. | .27 | 2.80** | .26 | 3.59** | 23 | 2.71** |
| Residential stability . ...... .. ... | .24 | 2 27* | −.12 | −1.74 | .11 | 1.40 |
| Immigrant concentration ...... .. | −.14 | −1.56 | −.22 | −3.51** | .12 | 1.52 |
| Population density ........ ........... | −.00 | −1.82 | − 35 | −5.10** | −.43 | −5 30** |
| Mixed land use . ......... ......... . | .00 | .12 | .03 | .51 | .08 | 1.06 |
| Collective efficacy ...... ....... ... ... | −1.97 | −4.12** | −.35 | −4.43** | −.27 | −2.92** |
| $R^2$ .......... ...... ......... ......... .. .... | 19 | | .46 | | .26 | |
| **Model 3:**[a] | | | | | | |
| SSO disorder ..... ..................... . | .09 | 1.56 | .10 | 2 12* | .02 | .29 |
| Collective efficacy ... ......... ... .. | −1.74 | −3.51** | − 11 | −2 38* | −.20 | −2.63** |
| Prior (1993) crime . ........... ...... | .26 | 2 02* | .83 | 21.01** | .59 | 10.15** |
| $R^2$ ... .................... . ........ ......... | 20 | | .85 | | 52 | |

NOTE.—$N = 195$ $R^2$ for homicide counts are pseudo $R^2$s; for log robbery rate and log burglary rate, they are adjusted $R^2$s. Homicide events are analyzed with negative binomial regression, with log population as the exposure (control) variable (coefficient not shown), robbery and burglary refer to logged incident rates per 100,000 (persons for robbery, households for burglary) analyzed with WLS regression using square root of population as weight For the 1995 homicide, robbery, and burglary equations, the control for prior crime refers to the 1993 log homicide count, robbery rate, and burglary rate, respectively

[a] All structural characteristics shown above are controlled (coefficients not shown)

* $P < 05$.

** $P < .01$.

on burglary (B = −.27). For present purposes, then, the key result is that the influences of structural characteristics and collective efficacy on burglary, robbery, and homicide are not mediated by neighborhood disorder.

As in table 4, we now exploit the longitudinal nature of the police data and provide a further test of collective efficacy and disorder by adding a control for the lagged crime rate (model 3). Unlike table 4, this specification turns out to be a severe one for official robbery and burglary because of the very strong connection between rates in 1993 and 1995, a result we suspect arises in part because of the common measurement source. For example, the 1993 log robbery rate is correlated .92 with the log robbery rate in 1995. This level of serial dependence makes it difficult for *any* concurrent factor to wield much direct influence, providing overly conser-

vative estimates.[25] Nevertheless, collective efficacy is directly related to lower rates of homicide, robbery, and burglary in 1995 after adjusting for 1993 levels of those crimes, as shown in the bottom portion of table 5. The coefficient for homicide ($-1.74$) remains large in magnitude, and the estimated effect for robbery ($t$-ratio $= -2.38$) is noteworthy in light of the very strong ($> .9$) positive association between 1993 and 1995 robbery coupled with the significant negative association between prior robbery and collective efficacy ($r = -.39$). By contrast, the association of disorder with homicide and burglary is reduced even further and is rendered insignificant for homicide.

The exception to the emerging conclusion that disorder is spuriously related to predatory crime is robbery. Columns 3–4 show that the effect of disorder is reduced but not eliminated by the introduction of a control for robbery in 1993. Areas with greater cues of disorder appear to be more attractive targets for robbery offenders, perhaps because disorder increases the potential pool of victims without full recourse to police protection, such as those involved in drug trafficking and prostitution. Wright and Decker's (1997) research has indicated that robbery offenders are especially attuned to local drug markets, where they perceive drug dealers and their customers as prime targets with cash on hand.[26]

Reciprocal Feedback

We close by addressing a concern that the estimated effect of collective efficacy on crime reflects reverse causation. It may be that neighborhood social trust and residents' sense of control are simultaneously undermined by crime, most notably interpersonal crimes of violence and those committed in public by strangers (Skogan 1990). Liska and Bellair's (1995) findings indicated that violent crimes such as robbery induce out-migration

[25] Note also that if collective efficacy has an insignificant direct effect in such a specification, this does not necessarily mean it is unimportant. Prior (unmeasured) levels of collective efficacy may have explained variation in prior crime, and thus the relationship with concurrent crime would be mediated.

[26] The three regressions of official crime rates were repeated with perceived disorder controlled; the main results were unchanged. The burglary and robbery models were also re-estimated using a negative binomial regression of raw counts, with logged population and logged counts of prior burglary and robbery controlled. The significant negative effect of collective efficacy on robbery and burglary was maintained, whereas the estimated effect of observed disorder on burglary was insignificant. Specifically, the negative binomial $t$-ratios reflecting the effect of collective efficacy on robbery and burglary events—with perceived disorder controlled—were $-2.95$ and $-2.57$, respectively. Further tests also revealed that physical disorder was somewhat more strongly related to robbery rates than social disorder when examined separately, presumably because physical disorder is the more reliable measure (see the appendix).

from central cities. Perhaps more relevant, Liska and Warner (1991) found that robbery constrains social interactions in public settings, thereby potentially dampening social cohesion and the emergence of shared expectations among residents for taking action to protect the community. Fear of crime, especially the fear of being accosted by strangers in public and attacked, may thus undermine neighborhood collective efficacy. By not accounting for the potential reciprocal effects of violence, we may have misestimated the role of collective efficacy relative to disorder. Our strategy to address this possibility has been to control for the prior incidence of violent events. Although providing a strict test, this strategy leaves unresolved the potential simultaneous relationship between collective efficacy and street crime. Moreover, because of the strong temporal dependence in the police record data, this strategy may have yielded unduly conservative estimates of both collective efficacy *and* disorder. We therefore address these issues by estimating a simultaneous equation model of violent crime and collective efficacy, and by introducing a new source of data on prior violence that is independent of police records.[27]

It is well known that estimating simultaneous relationships requires a priori identification restrictions that are, in practice, difficult to meet (Fisher and Nagin 1978). With respect to the issue at hand, however, we believe that extant theory provides reasonable grounds for specifying a causal feedback loop between violent crime and collective efficacy. First, we introduce measures of local social exchange, local friend/kinship ties, and neighborhood attachment as instrumental variables to identify the unique effect of collective efficacy.[28] A long line of urban research suggests that participation in social exchange, friend/kinship ties, and affective identification with the local area increases mutual trust and shared expectations for collective action in support of the neighborhood (Kasarda and Janowitz 1974; Sampson 1988). In the present case, we assume that friend/kinship ties, exchange, and affective sentiment for the neighborhood influence crime only insofar as they foster the core ingredients of collective efficacy. This seems reasonable, for there is little reason to expect that neighborly social ties or sentiment alone will reduce crime other than by

---

[27] We thank the reviewers for stressing the possibility of reciprocal causation, prompting us to push the analysis in new directions.

[28] The *local exchange* scale is based on five items in the survey that asked how often the respondent and neighbors exchange favors or goods (e.g., tools), have parties together, visit in each others' homes, watch each others' homes, and exchange advice or information. The measure of *friend/kinship ties* is derived from questions on the number of friends and relatives that respondents reported living in the neighborhood. *Attachment* is a two-item scale derived from questions on how sorry the respondent would be to move from the neighborhood and how satisfied he/she was with the neighborhood.

influencing the shared willingness to engage in informal social control under conditions of mutual trust (see also Bellair 1997).[29]

Second, to identify the crime equation, we created a new person-based index of the prevalence of prior homicide victimization among neighborhood residents, even if the incidents occurred elsewhere. The original source was death-record information (derived from the coroner's report) found in case-level vital statistics data for Chicago. Individual deaths by homicide victimization were geocoded and aggregated across years to yield more stable indicators. Based in part on a larger health-related project (Morenoff 1999), one indicator was constructed to reflect the age and sex-standardized rate of homicide victimization for the years 1989–91 for the larger NCs within which each tract is located. The second indicator is the logged rate of homicide deaths in each tract per 100,000 population for the years 1990–93. As expected, these two rates were positively correlated ($r = .62$; $P < .01$). To achieve parsimony, we combined them into a single index of resident-based homicide prevalence, which we use as an instrument to predict later incidents of violent crime.

Such a specification serves several theoretical purposes. It is well known in the criminological literature that victims and offenders in homicide transactions share similar demographic and residential profiles (Singer 1981). In fact, some 25% of homicides are victim precipitated (Wolfgang 1958). Thus, the prior rate of victim-based homicide is a proxy measure for resident offender rates of homicide production.[30] Moreover, it is plausible to assume that the homicide victimization rate of residents in 1989–93 will predict the later incidence of violence, but not directly affect levels of collective efficacy in 1995 independent of the concurrent rate of violent incidents. As Liska and Warner's (1991) research implies, it is the current presence of violence in their home neighborhoods—especially robbery—that residents most fear, not the past presence of victims, many of whom were victimized in other parts of the city. And, as noted above, there is evidence that residents are able to make the distinction between local

[29] Because local exchange and friend/kinship ties were rather highly correlated (.62; $P < .01$) and our goal is not to elucidate their independent effects but rather to identify the collective efficacy equation, for simplicity, we present the results for a combined scale of ties/exchange. We estimated models with each separately and found similar overall results.

[30] More specifically, this specification provides some purchase on assessing a "routine activity" model of violent events in the neighborhood while controlling for the differential composition of neighborhoods with respect to residents' involvement with violence. We acknowledge an anonymous reviewer for suggesting that we consider the violence potential of neighborhood residents. We sought to obtain offender residence data but were unsuccessful. The Chicago police do not release this information, and it is not available in vital statistics. Therefore we cannot directly assess offender production.

crime events and persons involved in criminal networks (Pattillo 1998). Therefore, we assume that any effect of resident-based homicide prevalence in 1989–93 on collective efficacy in 1995 works through its connection to the 1995 event rate of violence.

The maximum-likelihood estimates of the coefficients and $t$-ratios for the logged rate of homicide incidents per capita in 1995 are presented in table 6.[31] The fit of the model to the data is very good, with a chi-square of 9.86 relative to 6 degrees of freedom ($P > .10$) and an adjusted goodness of fit index of .90. The coefficients and standard errors in table 6 reveal that the estimated direct effects of the instrumental variables are substantial. Social ties/exchange and attachment each predict collective efficacy net of structural controls and homicide, while the direct effect of prior homicide victimization on the 1995 homicide rate is also significant ($t$-ratio = 2.85). The instrumental variables accounted for 22% of the unique variance in collective efficacy and 34% of the unique variance in homicide rates in the reduced-form equations. Most important, the fitted structural model indicates the presence of a reciprocal feedback between homicide and collective efficacy. Note the $t$-ratio of $-2.20$ for the estimated effect of collective efficacy on homicide and the $t$-ratio of $-2.58$ for the estimated effect of homicide on collective efficacy. Controlling for this feedback loop does not change our inference about the direct effect of social disorder on homicide. It remains insignificant ($t$-ratio = $-.38$).

Prior research on the fear of crime (Liska and Warner 1991; Liska and Bellair 1995) suggests that robbery might present an even more critical test, and our own data has shown that robbery is the only crime directly linked to social disorder (table 5). We thus estimated a simultaneous equation model using robbery instead of homicide as the measure of violence (not shown here, details available upon request). The results were in most respects similar to those in table 6. The instrumental variable for robbery was surprisingly effective, with prior homicide accounting for 30% of the unique variance in 1995 robbery.[32] And once again, there was a reciprocal association between collective efficacy and violence, with collective efficacy negatively related to robbery ($t = -2.40$) and robbery related nega-

---

[31] The models were estimated with LISREL ver. 8.04 (Joreskög and Sörbom 1993), with a weighted full-information, maximum-likelihood (FIML) procedure (unweighted results were similar). Recall that the logged-rate and event-count models produced similar results for both homicide and robbery (table 5). Also, recall that mixed land use consistently showed a large association with SSO disorder but not crime rates (see tables 4–5). Further analysis revealed an insignificant association of land use with collective efficacy. We therefore excluded land use in both the crime rate and collective efficacy equations, improving the overall model fit.

[32] The fit of the robbery model was not quite as good as for homicide, but it was still very adequate ($\chi^2 = 15.2$; $df = 6$; $P = 02$, AGFI = 85).

## TABLE 6

MAXIMUM-LIKELIHOOD COEFFICIENTS AND $t$-RATIOS FOR SIMULTANEOUS EQUATION MODEL OF SSO DISORDER, COLLECTIVE EFFICACY, AND OFFICIAL HOMICIDE RATES

| | ENDOGENOUS VARIABLES | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | Collective Efficacy[a] | | SSO Disorder | | Homicide Rate[a] | |
| | Coefficient | $t$-ratio | Coefficient | $t$-ratio | Coefficient | $t$-ratio |
| Exogenous: | | | | | | |
| Concentrated disadvantage ............ | −.04 | −2.26* | .89 | 7.74** | −.04 | −.17 |
| Residential stability .... ............... | .08 | 4.93** | −.02 | −.17 | .44 | 2.14* |
| Immigrant concentration .. .. .. ... .... | −.04 | −3.61** | .51 | 4.64** | −.11 | −.62 |
| Population density . ......... ... ....... | −.00 | −2.83** | .00 | .27 | .00 | .15 |
| Mixed land use  ... ...... .. ... .... ... ... | NI | | .04 | 4.19** | NI | |
| Ties/exchange[b] ......... .. .. ... .. .... | .03 | 3.91** | NI | | NI | |
| Attachment[b] ... ..... .. .. ... ... ... ... | .11 | 3.25** | NI | | NI | |
| Prior homicide[b] ............... ... ... ... | NI | | NI | | .37 | 2.85** |
| Endogenous. | | | | | | |
| Collective efficacy ............  ... .. | NI | | −1.69 | −2.64** | −3.18 | −2.20† |
| SSO disorder ............ ... .. .... | NI | | NI | | −.03 | −.38 |
| Homicide rate  ...... ............ ... .. . | −.06 | −2.58* | NI | | NI | |
| $R^2$ ......................................... | .63 | | .58 | | .32 | |

NOTE.—$N = 195$  For overall fit, $\chi^2 = 9.86$; $df = 6$; and $P > 10$. NI = Not included in model specification
[a] Error covariance for reciprocal relationship between collective efficacy and homicide = 8.72 ($t$-ratio = 2.24)
[b] Instrumental variable.
* $P < .05$.
** $P < .01$.
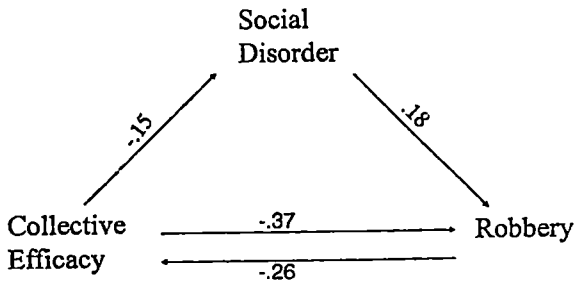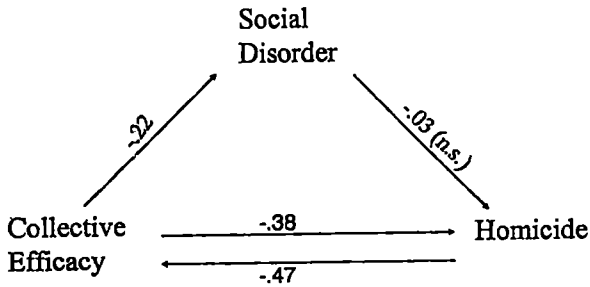
FIG. 1.—Structural model for relationships among endogenous variables—survey reported collective efficacy, observed social disorder, and official rates of violence. Coefficient estimates are from standardized solution and significant except where noted. For simplicity, exogenous variables are not displayed (see table 6 for a full listing of coefficients and *t*-ratios for homicide).

tively to collective efficacy ($t = -2.10$). Controlling for this reciprocal relationship, there remained a significant direct effect of social disorder on robbery ($t$-ratio $= 2.68$).

To aid in interpreting the pattern and magnitude of these results across crime types, standardized coefficients for the key structural relations of interest are shown in figure 1. The reciprocal relationship between collective efficacy and homicide is seen in the top half, with a collective efficacy to homicide path of $-.38$ and a homicide to collective efficacy path of $-.47$. The bottom half of figure 1 shows that robbery and collective efficacy are caught up in a similar feedback loop, net of structural characteristics, with a standardized coefficient of $-.37$ for the estimated effect of collective efficacy on robbery and $-.26$ for the reverse effect of robbery on collective efficacy. Interestingly, further comparisons reveal that for

635

both robbery and homicide, the unstandardized coefficients for collective efficacy change very little in magnitude when the simultaneous relationship is modeled. For example, the coefficient estimate for collective efficacy on robbery in a recursive specification is $-1.29$ (SE $= .29$), compared to $-1.33$ (SE $= .55$) for the simultaneous model in the bottom half of figure 1. The corresponding coefficients for homicide are $-3.54$ (SE $= .78$) and $-3.18$ (SE $= 1.57$). Thus, the standard errors increase in the simultaneous equation results, but the substantive pattern remains the same (cf. tables 4–5).

The other major result obtained in figure 1 is that collective efficacy is inversely related to SSO disorder, which in turn exhibits a significant positive association with the robbery rate.[33] The standardized coefficient for SSO disorder on robbery is B $= .18$, similar in magnitude to the negative effect of collective efficacy on disorder (B $= -.15$). By contrast, the estimated effect of disorder on homicide is virtually zero. These patterns are consistent with earlier results from recursive models. To assess robustness, we re-estimated the models in figure 1 by specifying ties/exchange as the sole instrument for collective efficacy, and by separately introducing the constituent lagged indicators of homicide victimization as instruments for the current event rate. The results were substantively similar. We also controlled for the effects of perceived disorder on collective efficacy in both the robbery and homicide equations, specifying the former as a function of observed disorder. The results indicated that observed disorder increases perceived disorder, which in turn reduces collective efficacy. The significant reciprocal relationship between violence and collective efficacy nonetheless remained intact, as did the differential effect of observed disorder in predicting robbery but not homicide in figure 1.[34] The overall pattern in the data is thus clear. The estimated association of collective efficacy with disorder and crime holds up for a variety of measures and alternative theoretical specifications, including simultaneous causation, and thereby appears to be general in nature. By contrast, the direct disorder-crime association is specific to officially measured robbery.

[33] The estimated indirect effect of collective efficacy on robbery mediated by SSO disorder for the model in the bottom half of figure 1 is significant at $P < .10$ ($t$-ratio $= -1.62$).

[34] Furthermore, we estimated a model specifying a simultaneous relationship between collective efficacy and SSO disorder, using mixed land use as an instrument for disorder and, as in table 6, ties/exchange and attachment as instruments for collective efficacy. Yielding an excellent fit to the data ($\chi^2 = .005$; $df = 1$, $P = .94$), the model results indicated that the path from observed disorder to collective efficacy was insignificant, whereas the $t$-ratio for the estimated effect of collective efficacy on observed disorder was $-2.74$ ($P < .01$).

SUMMARY

The evidence presented in this study has shown that public disorder in urban spaces is a robust ecological construct that can be reliably measured at the neighborhood level using systematic observational procedures. Consistent with our theoretical expectations, structural characteristics—especially concentrated poverty and mixed land use—were strongly associated with physical and social disorder. Yet collective efficacy, defined as the fusion of social cohesion with shared expectations for the active social control of public space, predicted lower observed disorder after controlling not just sociodemographic and land-use characteristics, but perceived disorder and prior rates of predatory crime as well. Collective efficacy also maintained a significant relationship with violent crime after adjusting for simultaneous feedback effects.

On the other hand, observed disorder did not match the theoretical expectations set up by the main thesis of "broken windows" (Wilson and Kelling 1982; Kelling and Coles 1996). Disorder is a moderate correlate of predatory crime, and it varies consistently with antecedent neighborhood characteristics. Once these characteristics were taken into account, however, the connection between disorder and crime vanished in 4 out of 5 tests—including homicide, arguably our best measure of violence. The empirical results therefore support our contention that public disorder and most predatory crimes share similar theoretical features and are consequently explained by the same constructs at the neighborhood level, in particular the concentration of disadvantage and lowered collective efficacy.

Although our results contradict the strong version of the broken windows thesis, they do not imply the theoretical irrelevance of disorder. After all, our theoretical framework rests on the notion that physical and social disorder comprise highly visible cues to which neighborhood observers respond (see also Jacobs 1961; Goffman 1963; Lofland 1973; Skogan 1990; Taylor 1997). According to this view, disorder may turn out to be important for understanding migration patterns, investment by businesses, and overall neighborhood viability. Thus, if disorder operates in a cascading fashion—encouraging people to move (increasing residential instability) or discouraging efforts at building collective responses—it would indirectly have an effect on crime. Moreover, our results established a significant albeit relatively modest association of disorder with officially measured robbery. Apparently, robbery not only constrains social interaction and thus reduces social control (Liska and Warner 1991), but potential robbery offenders respond to visual cues of social and physical disorder in the neighborhood. Our findings regarding robbery also suggest a complex feedback loop (bottom half of figure 1), whereby disorder entices robbery,

which in turn undermines collective efficacy, leading over time to yet more disorder and ultimately robbery.

What we would claim, however, is that the current fascination in policy circles (see Kelling and Coles 1996; Kelling 1998) on cleaning up disorder through law enforcement techniques appears simplistic and largely misplaced, at least in terms of directly fighting crime. Eradicating disorder *may* indirectly reduce crime by stabilizing neighborhoods, but the direct link as formulated by proponents was not the predominate one in our study. What we found instead is that neighborhoods high in disorder do not have higher crime rates in general than neighborhoods low in disorder once collective efficacy and structural antecedents are held constant (tables 4–5). Crime and disorder are not even that highly correlated in the first place. Even for robbery, the aggregate-level correlation does not exceed .5.[35] In this sense, and bearing in mind the example of some European and American cities (e.g., Amsterdam, San Francisco) where visible street-level activity linked to prostitution, drug use, and panhandling does not necessarily translate into high rates of violence, public disorder may not be so "criminogenic" after all in certain neighborhood and social contexts (see also Whyte 1988, pp. 156–64). Put differently, the active ingredients in crime seem to be structural disadvantage and attenuated collective efficacy more so than disorder. Attacking public disorder through tough police tactics may thus be a politically popular but perhaps analytically weak strategy to reduce crime, mainly because such a strategy leaves the common origins of both, but especially the last, untouched. A more subtle approach suggested by this article would look to how informal but collective efforts among residents to stem disorder may provide unanticipated benefits for increasing collective efficacy (Skogan and Hartnett 1998), in the long run lowering crime.[36]

Of course, several limitations of our study warrant further consideration. First, although the relationships that emerged were consistent under multiple and strict tests, until we have a longitudinal profile of neighbor-

---

[35] We also set aside for future research the possibility that the robbery finding reflects an artifact of official data. Recall that none of the survey measures of violence were directly related to disorder. It is possible that citizen calls to the police or police accuracy in recording robberies is greater in areas perceived to be high in disorder. That there was no disorder link for official homicide or burglary is perhaps telling.

[36] Informally mobilizing a neighborhood "clean up" to reduce physical disorder, for example, might build collective efficacy through the formation of new social ties and by increasing local awareness of the mutual commitment of residents to the area. Such mobilization might also demonstrate to participants and observers alike that people in the neighborhood could be relied upon to maintain public order and social control. A demonstration of formal control through a police-led crackdown on disorder might be expected to generate a different response, unless it too was internally mobilized.

hood change, the causal direction of effects rests in large part on a priori theory. Understanding the dynamics of disorder is a complex challenge and will require a new generation of longitudinal research. Second, and relatedly, our simultaneous equation model must be viewed cautiously despite its apparent robustness. Definitive tests of feedback processes await future research and validation. Third, the SSO method is limited to the extent that inferences from the visual record were systematically biased in ways not anticipated by the investigative team. It is possible, for example, that the trained observers used physical cues of disorder to inform judgments about social disorder.[37] The SSO training tried hard to avoid this scenario, but the fact is that all observations require inferences. Fourth, we acknowledge that the large number of ecological units assessed means that many contextual features were probably missed. SSO is thus not meant to be a replacement for the close ethnographic observation of neighborhoods found in a long-standing and continuing tradition of research (e.g., Whyte 1943; Lofland 1973; Pattillo 1998; Carr 1998). Rather, the whole point of *systematic* social observation is to allow for the comparative analysis of variations across a large number of analytically defined ecological areas. Our goal has therefore been to strive for a systematic method that is replicable (see also King, Keohane, and Verba 1993) and that serves as a complement to comparative urban ethnography. A fifth limitation to the approach proposed in this article is that we have not yet taken into account spatial autocorrelation (Raudenbush and Sampson 1999b). Rather, tract-level neighborhoods have been treated as independent.[38] Ongoing work will build spatial associations into the types of models presented here.

IMPLICATIONS FOR RESEARCH DESIGN

This article is part of a larger effort to build a *social* science of ecological assessment by formulating new research designs and statistical methods to improve the quality of "ecometric" measures (see also Raudenbush and Sampson 1999a, 1999b). Whereas psychometric procedures for the study

[37] However, this possibility is undermined by the fact that the results were not sensitive to substitution of the physical disorder for the social disorder SSO measure.

[38] This is not an unreasonable assumption for the SSO design because many of the analyzed tracts ($N = 195$) are not contiguous to other sampled tracts. It is thus not clear what a spatial model would accomplish when less than 25% of the total tracts in the city ($N = 865$) are studied, effectively censoring the majority of contiguous spatial units. Nevertheless, modeling spatial dependence between tracts might reduce "noise" introduced by spatially correlated errors Information about spatial dependence might also make it possible to obtain reasonable measures of neighborhood disorder even for areas sparsely assessed by direct observation.

of individual-level properties are well established in the behavioral sciences, the development of ecometric research procedures and statistical methods for the study of ecological and other macrolevel units is in its infancy. We have explored the systematic social observation (SSO) of public spaces linked to neighborhood surveys, census data, and police records in an effort to create reliable and valid measures of neighborhood-level disorder. The SSO is an especially important case for "ecometrics" in sociology given the potential utility of videotaping as an observational strategy in the study of neighborhoods and other collectivities.

Consider some of the benefits for neighborhood research designs. Unlike on-site assessments, videotapes can be stored and made available for future researchers. They can be retrieved and revisited on demand, whether for assessing interrater reliability, for recoding, or to construct new variables not considered by the original investigators. An even richer possibility concerns the generative potential of video. Suppose that one finds an unusual concentration of criminal activity in certain neighborhoods that cannot be easily explained on the basis of measurable demographic characteristics or social processes like collective efficacy. Researchers can go back to the videos for selected face blocks to look anew at what might lie behind the density of criminal activity. Although not used as such in the present article, this possibility for video-based SSO is similar to the use of ethnography to generate hypotheses for future inquiry.[39]

Another advantage pertains to the multilevel assessment of neighborhood effects in sociological and developmental research. One of the central problems of extant research is that neighborhoods are treated as static constructs even though we know that neighborhoods change, often rapidly. Moreover, individuals move frequently during the course of longitudinal studies, yet typically, census tract data are "assigned" to individuals based on their past residence. Videotaping offers a relatively cheap and effective way to track within-neighborhood change and changes in the neighborhoods of residence. For example, as individuals move, interviewers at follow-up contacts could adopt an SSO strategy while in the process of carrying out individual-level assessments. Videotape records could then be matched to prior tapes and coded for change. When integrated with census data and perhaps surveys, the possibilities for discovering the ways in which neighborhoods influence individual development are greatly enhanced. Moreover, because face blocks are the first unit of measurement, "neighborhoods" can later be defined by the researcher at varying levels

---

[39] Beyond the scope of this article are potential ethical dilemmas in the use of systematic videotaping procedures, such as privacy and informed consent by street participants ("human subjects"?), the unique identification of individuals, and the appropriation of videotapes by the police for use as evidence of crimes.

of aggregation. Based on theory or emergent findings, for example, adjacent face blocks might be pieced together to form ecological units that better conform to the processes at hand. In this way, the use of SSO at the microlevel of blocks offers maximum flexibility for neighborhood-level research.

Our techniques for SSO also forge an explicit link with the technological advances that are transforming the ways in which research is conducted. Advertisers, fund-raisers, and market researchers are far ahead of sociologists in their command of geographic databases to cull sociological information and construct profiles of community process. With face blocks as the unit, SSO video can be linked to geographic information system (GIS) databases that allow for instantaneous merger with rich sources of information. For instance, SSO could in principle be linked to address-level databases on employment, real estate sales, and building code violations.

Finally, and perhaps most important, SSO provides the sights, sounds, and feel of the streets. Much as practiced by the original Chicago school of urban sociology (Abbott 1997), SSO takes researchers to the streets in a very real way. The present article has only scratched the surface of the potential for such a take on community. We thus believe the final story line pertains not just to collective efficacy or public disorder as theoretical constructs, but the potential scope of systematic social observation as an analytic tool. Visual cues are salient in many dimensions of social life; systematically observing them in their natural social context should be, as Reiss (1971) and Whyte (1988) argued, and a generation of students of the city before them (e.g., Park and Burgess 1921), a fundamental part of the sociological enterprise.

## APPENDIX

### A Model for Uncertainty in Systematic Social Observation

To estimate the magnitude of measurement errors in scales derived from SSO at the NC ($N = 80$) level, Raudenbush and Sampson (1999$b$) formulated a three-level model. We adapt that model to the data at hand, with multiple items nested within face blocks nested within census tracts. Level 1 captures the error due to item inconsistency within a face block; level 2 captures error due to variation between face blocks within census tracts while also adjusting for time-of-day effects; and level 3 specifies the "true-score" variation, that is, the variation between census tracts ($N = 196$) on the latent variable of interest. The model produces reliability estimates and the estimated correlation between latent variables adjusting for measurement error at each level. Because the item responses are dichotomous (presence or absence of each indicator of disorder within a face block), the model is a three-level logistic regression model.

Let $Y_{ijk}$ take on a value of unity if indicator $i$ of disorder is found present in face block $j$ of tract $k$, with $Y_{ijk} = 0$ if not; and let $\mu_{ijk}$ denote the probability $Y_{ijk} = 1$. As is standard in logistic regression, we define $\eta_{ijk}$ as the log-odds of this probability. Thus, we have

$$Y_{ijk}|\mu_{ijk} \sim \text{Bernoulli};$$

$$E(Y_{ijk}|\mu_{ijk}) = \mu_{ijk}$$

$$Var(Y_{ijk}|\mu_{ijk}) = \mu_{ijk}(1 - \mu_{ijk}).$$

$$\eta_{ijk} = \log\left(\frac{\mu_{ijk}}{1 - \mu_{ijk}}\right).$$

At level 1, the log-odds of finding disorder on item $i$ depends on which aspect of disorder is of interest (physical or social) and which specific item is involved. Let $D_{pijk}$ take on a value of 1 if item $i$ is an indicator of physical disorder, 0 otherwise; and let $D_{sijk} = 1$ similarly take on a value of 1 if that item indicates social disorder. Then we have

$$\eta_{ijk} = D_{pijk}\left(\pi_{pjk} + \sum_{m=1}^{9} \alpha_{mjk}X_{mijk}\right) + D_{sijk}\left(\pi_{sjk} + \sum_{m=1}^{6} \delta_{mjk}Z_{mijk}\right),$$

where $X_{mijk}$, $m = 1, \ldots, 9$ are dummy variables representing nine of the ten items that measure physical disorder (each taking on a value of 1 or 0); $Z_{mijk}$, $m = 1, \ldots, 6$ are dummy variables representing six of the seven items that measure social disorder. We "center" each $X$ and $Z$ around its grand mean, which leads to the following definitions: $\pi_{pjk}$ is the adjusted log-odds of finding physical disorder on a "typical item" when observing face block $j$ of tract $k$; $\pi_{sjk}$ is the adjusted log-odds of finding social disorder on a "typical item" when observing face block $j$ of tract $k$; $\alpha_{mjk}$ reflects the "severity" level of item $m$ within the physical disorder scale;[40] similarly, $\delta_{mjk}$ reflects the "severity" level of item $m$ within the social disorder scale.[41] The item difficulties alpha and delta could, in principle, be allowed to vary across face blocks or NCs; however, in the absence of theory that

[40] Interpreting these coefficients in terms of "severity" requires that they be multiplied by $-1$ (Raudenbush and Sampson 1999b).

[41] One benefit of this model is that face-block measures of disorder, $\pi_{pjk}$ and $\pi_{sjk}$, are adjusted for missing data. Because of the expense of coding the videotapes, we coded a random subsample of face blocks within tracts. Face blocks not sampled have data from the observation log but not the coding log. No bias arises because the coded face blocks constituted a representative sample of face blocks in the tracts. Nevertheless, controlling the item difficulties enables all of the data to be effectively used in the analysis (Raudenbush and Sampson 1999b).

might predict such variation, they will be held constant in the interest of parsimony. Thus, $\alpha_{mjk} = \alpha_m$ and $\delta_{mjk} = \delta_m$ for all $j$, $k$.

The level-2 model accounts for variation between face blocks within tracts on latent face-block disorder. Each is predicted by the overall tract-level disorder and the time of day during which the face block was observed:

$$\pi_{pjk} = \beta_{pk} + \sum_{q=1}^{5} \theta_{pqk}(\text{time})_{qjk} = u_{pjk},$$

$$\pi_{sjk} = \beta_{sk} + \sum_{q=1}^{5} \theta_{sqk}(\text{time})_{qjk} = u_{sjk},$$

$(\text{time})_{qjk}$ for $q = 1, \ldots, 5$ are five time-of-day indicators. They indicate the hours 7:00 to 8:59 A.M.; 9:00 to 10:59 A.M.; 11:00 A.M. to 12:59 P.M.; 1:00 to 2:59 P.M.; and 3:00 to 4:59 P.M., where the omitted group is from 5:00 to 6:59 P.M.; $\theta_{pqk}$ and $\theta_{sqk}$ are regression coefficients that capture the time-of-day effects on observing physical and social disorder within tract $k$. These could be allowed to vary over tracts, but for parsimony we hold them constant such that $\theta_{pqk} = \theta_{pq}$ and $\theta_{sqk} = \theta_{sq}$ for all $k$. Here $\beta_{pk}$ and $\beta_{sk}$ are the "true" scores for tract $k$ on physical and social disorder, respectively, adjusting for time of day. The random effects $u_{pjk}$, $u_{sjk}$ are assumed bivariate normally distributed with zero means, variances $\tau_{pp}$ and $\tau_{ss}$, and covariance $\tau_{ps}$. The variances will be large when face blocks vary greatly within tracts on disorder.

The third and final level of the model describes variation between tracts, the key units of neighborhood measurement, on physical and social disorder. We have simply

$$\beta_{pk} = \gamma_p + \upsilon_{pk},$$

$$\beta_{sk} = \gamma_s + \upsilon_{sk},$$

where $\gamma_p$ and $\gamma_s$ are the grand mean levels of physical and social disorder in Chicago neighborhoods and the random effects $\upsilon_{pk}$ and $\upsilon_{sk}$ are assumed bivariate normally distributed with zero means, variances $\omega_{pp}$ and $\omega_{ss}$, and covariance $\omega_{ps}$. The variances will be large when tracts vary greatly on their levels of disorder.

## Measurement Results

The three models were combined as described in Raudenbush and Sampson (1999b), and all model parameters were estimated simultaneously by penalized quasi likelihood (Breslow and Clayton 1993) using an algorithm

TABLE A1

LEVEL-1 HLM RESULTS FOR VARIATION IN SSO DISORDER ITEMS
WITHIN FACE BLOCKS: ITEM DIFFICULTY

| Item | Coefficient | SE |
|---|---|---|
| Physical disorder: | | |
| Intercept .. ..... ................. .. . .. .. .. ......... | −2.30** | 10 |
| Cigarettes, cigars in street/gutter .. ......... | 3.36** | .03 |
| Garbage/litter on street/sidewalk .. .. ..... | 2.33** | .03 |
| Empty beer bottles visible in street .. .. ..... | 1.11** | .03 |
| Tagging graffiti ..... ...... . .. . .................. .... | .33** | 04 |
| Graffiti painted over ......... .. .. . . ........... | (reference item) | |
| Gang graffiti ... .. .. .. ................ .. . . . .. | −.68** | 05 |
| Abandoned cars . .. ...................... .. . .. .. .. . | −1.22** | 05 |
| Condoms on sidewalk . . ................. .... | −2.63** | .08 |
| Needles and syringes . . . ..... .. .. ............. | −2.90** | .09 |
| Political message graffiti . . .. . .. .. ......... | −5 08** | .29 |
| Social disorder | | |
| Intercept . .. .. ....... .. . . . .................. .. .. | −6 96** | 19 |
| Adults loitering or congregating ............... | 3 83** | 24 |
| People drinking alcohol . . .. .. . ........... | 61* | .29 |
| Peer group, gang indicators present ... .. ..... | (reference item) | |
| People intoxicated .. ...................... .. .. . ..... | −.06 | .34 |
| Adults fighting or hostilely arguing .......... | −.41 | .37 |
| Prostitutes on street . . .... . ..................... ... | −.49 | .38 |
| People selling drugs ... .. . .. . . .. . .. ....... ... | −.81 | .42 |

* P < 05
** P < 01

described in detail by Raudenbush (1995) and now implemented in version 4 of the HLM program (Bryk et al. 1996). Table A1 presents the estimates for the item response model. Items with negative coefficients have low probabilities of occurrence and thereby are rarer and more "severe" than are items with positive coefficients. In the physical disorder scale, the presence of cigarettes or cigars on the street or sidewalk, garbage and litter, along with the presence of empty beer bottles are comparatively less severe than the presence of gang graffiti, abandoned cars, or condoms on the sidewalk. Item severity thus conforms to intuitive expectations. The exception is political graffiti, which is exceptionally rare yet not generally regarded as especially severe. The item severities for physical disorder vary substantially, a feature of a well-behaved scale. In contrast, the item severities in the social disorder scale are clumped at the severe end except for adults loitering or congregating and drinking. This pattern reflects the low frequency of the social disorder indicators. Although the item severities are not well separated, their ordering does correspond to theoretical

TABLE A2

LEVEL-2 HLM RESULTS FOR VARIATION IN SSO
DISORDER ITEMS BETWEEN FACE BLOCKS WITHIN
TRACTS: TIME OF DAY EFFECTS

| Time | Coefficient | SE |
|---|---|---|
| Physical disorder: | | |
| 7:00–8 59 ..... ... . . ...... | 25** | 05 |
| 9:00–10.59 .. . ........... | .09** | .03 |
| 11·00–12:59 . .......... . . | .10* | .04 |
| 1:00–2 59 .......... .. ...... | .15** | .04 |
| 3:00–4:59 ....... .. .. ........ | 05 | 04 |
| 5·00–6:59 .... . .... ........ . | (reference time) | |
| Social disorder | | |
| 7.00–8:59 .. ........... .... . | − 72** | .20 |
| 9:00–10:59 .... . . ........ . | − 80** | .13 |
| 11·00–12.59 . .. ............. . | −.53** | 17 |
| 1·00–2:59 .. . ........... . | −.00 | .14 |
| 3.00–4·59 ..... ........ .. ... | − 09 | .12 |
| 5:00–6·59 ............... . .. .. | (reference time) | |

* P < 05
** P < 01

expectation, with adults loitering and drinking alcohol being less severe
than adults fighting, prostitution, or drug sales.

Table A2 provides estimates of the effects of time of day. Presumably
social interactions in public view occur with relatively little frequency
early in the morning and more frequently later on. This appears to be
true of those social interactions involving some element of disorder as
well. Note the mainly positive trend in time for social disorder, with coef-
ficients of −.72, −.80, −.53, −.00, and −.09 as the day progresses. As
expected, no such trend is apparent in the case of physical disorder. All
model estimates in this article are adjusted for time-of-day effects.

Variance-Covariance Components and Reliability

The estimation of the variance-covariance components in table A3 pro-
vides information that we use to assess the quality of SSO measures. For
comparative purposes, we report results for both the 196 census tracts
and the larger NCs ($N = 80$). The intratract correlations are .46 and .98
for physical and social disorder, respectively. Formally,

$$\rho_{TRp} = \frac{\omega_{pp}}{\omega_{pp} + \tau_{pp}}; \; \rho_{TRs} = \frac{\omega_{ss}}{\omega_{ss} + \tau_{ss}}.$$

TABLE A3

LEVEL-3 HLM RESULTS: VARIANCE-COVARIANCE COMPONENTS
AND MEASUREMENT PROPERTIES FOR DISORDER SCALES AT FACE
BLOCK AND TRACT LEVELS

|  | Physical Disorder | Social Disorder |
|---|---|---|
| Between face blocks: |  |  |
| Variance ....... ... .. ... ........... .. ........ | .67** | .02 |
|  | (.02) | (.16) |
| Covariance ........ .. .. ............... ...... | .06 |  |
|  | (.05) |  |
| Reliability .. .. ............ .... . .. ........... | .37 | .00 |
| Between census tracts: |  |  |
| Variance ......... ..... ............. .. . ...... | .56** | 1 12** |
|  | (.06) | ( 16) |
| Covariance ........ . .. ..................... .. | .47** |  |
|  | ( 08) |  |
| Reliability .... . ............... . .......... | 95 | .72 |

\* *P* < .05
\*\* *P* < 01.

Thus, the intratract correlations, $\rho_{TRp}$ and $\rho_{TRs}$ for physical and social disorder, respectively, are the ratios of variance between tracts to overall variance, after adjusting for item inconsistency. The intra-NC correlation for physical disorder is .39 (Raudenbush and Sampson 1999*b*, pp. 28–29). The fact that the intratract correlation is larger than the intra-NC correlation arises because census tracts are internally more homogeneous than are NCs. The very high intratract correlation of .98 for social disorder suggests that there is little reliable variation between face blocks within tracts. Because evidence of social disorder is rare, data on social disorder at the face-block level are sparse, leading to imprecise estimation of the within-tract variance component.

Closely related to the intratract correlation is the internal consistency reliability of tract-level measurement. The latter depends on the intratract correlation but also on the number of face blocks sampled and the item severities. If all items are "severe," evidence of disorder is hard to find, and reliability will be comparatively low. Raudenbush and Sampson (1999*b*) formulate the reliability for tract *k,* in the case of physical disorder, as given by

$$\lambda_{pk} = \frac{\omega_{pp}}{\omega_{pp} + \dfrac{\tau_{pp}}{J_k} + \dfrac{1}{nJ_kw}},$$

where $\lambda_{pk}$ is the reliability of the physical disorder measure for tract $k$; $n = 10$ is the number of items in the scale; $J_k$ = the number of face blocks sampled in tract $k$; and $w$ = the average value of the estimates of $\mu_{yk}(1 - \mu_{yk})$ in tract $k$. The formula for social disorder is analogous, revealing that reliability will be high when (a) the between-tract variance $\omega_{pp}$ is large relative to the within-tract variance $\tau_{pp}$; (b) the number of items in the scale, $n$, is large; (c) the number of face blocks sampled, $J_k$, is large; and (d) the probability of finding an item of disorder in a given face block ($\mu_{yk}$) is near .50, at which point $w$ achieves its maximum.

The results in table A3 show that for physical disorder, the average of these reliabilities is .95. This extremely high reliability reflects a reasonably high intra–face-block correlation of .46, the large number of face blocks per tract, and the well-behaved item severities noted earlier. The average reliability for social disorder is lower than for physical disorder but still relatively high (.72). The difference reflects the extreme rarity of many of the indicators of social disorder and exists despite the high intra-tract correlation and the large number of face blocks per tract (98). Interestingly, the reliabilities for the 80 NCs studied by Raudenbush and Sampson (1999b) are not much different: .98 and .83 for physical disorder and social disorder, respectively. As tracts afford much greater power than NCs to detect between-neighborhood effects and are also more homogeneous internally, the similarities in reliabilities provide further evidence supporting the use of tracts as units of analysis. What about even lower levels of aggregation? Unfortunately, but not surprisingly given our results, reliabilities at the face-block level are simply unacceptable: .37 for physical disorder and .00 for social disorder.

REFERENCES

Abbott, Andrew. 1997. "Of Time and Space: The Contemporary Relevance of the Chicago School." *Social Forces* 75:1149–82.
Bellair, Paul. 1997. "Informal Surveillance and Street Crime: Disentangling Reciprocal Effects." Paper presented at the annual meeting of the American Sociological Association, Toronto, August
Belsley, David, Edwin Kuh, and Roy Welsch. 1980 *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York: Wiley.
Brantingham, Patricia, and Paul Brantingham. 1995. "Environment, Routine, and Situation: Toward a Pattern Theory of Crime." *Advances in Criminological Theory* 5:259–94.
Breslow, N., and D. G. Clayton. 1993 "Approximate Inference in Generalized Linear Mixed Models." *Journal of the American Statistical Association* 88:9–25.
Bryk, Anthony, and Stephen Raudenbush. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Newbury Park, Calif.: Sage.
Bryk, Anthony, Stephen Raudenbush, Richard Congdon, and M. Seltzer. 1996. *HLM: Hierarchical Linear and Nonlinear Modeling with the HLM/2L and HLM/3L Programs.* Chicago: Scientific Software International.

Bursik, Robert J., Jr. 1988. "Social Disorganization and Theories of Crime and Delinquency: Problems and Prospects." *Criminology* 26:519–52.

Bursik, Robert J , Jr., and Harold Grasmick. 1993. *Neighborhoods and Crime: The Dimensions of Effective Community Control.* New York: Lexington.

Carr, Patrick J. 1998. "Keeping Up Appearances: Informal Social Control in a White Working-Class Neighborhood in Chicago." Ph.D. dissertation. University of Chicago, Department of Sociology

Carter, Woody, Jody Dougherty, and Karen Grigorian. 1996 "Videotaping Neighborhoods." Working paper. University of Chicago, National Opinion Research Center.

Cohen, Lawrence, and Marcus Felson. 1979 "Social Change and Crime Rate Trends: A Routine Activity Approach." *American Sociological Review* 44:588–608.

Cohen, Lawrence, James Kluegel, and Kenneth Land. 1981. "Social Inequality and Predatory Criminal Victimization: An Exposition and Test of a Formal Theory." *American Sociological Review* 46:505–24.

Covington, Jeanette, and Ralph B. Taylor. 1991. "Fear of Crime in Urban Residential Neighborhoods. Implications of Between- and Within-Neighborhood Sources for Current Models." *Sociological Quarterly* 32:231–49.

Farrington, David 1979. "Longitudinal Research on Crime and Delinquency." Pp. 289–348 in *Crime and Justice,* vol. 1. Edited by Norval Morris and Michael Tonry. Chicago: University of Chicago Press

Felson, Marcus. 1987. "Routine Activities and Crime Prevention in the Developing Metropolis." *Criminology* 25:911–31.

Fisher, Franklin, and Daniel Nagin. 1978. "On the Feasibility of Identifying the Crime Function in a Simultaneous Model of Crime Rates and Sanction Levels." Pp. 361–99 in *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates,* edited by Alfred Blumstein, Jacqueline Cohen, and Daniel Nagin. Washington, D.C.. National Academy Press.

Garofalo, James, and John H. Laub 1978. "The Fear of Crime. Broadening Our Perspective." *Victimology* 3:242–53.

Goffman, Erving. 1963. *Behavior in Public Places: Notes on the Social Organization of Gatherings.* New York. Free Press.

Gottfredson, Michael, and Travis Hirschi. 1990. *A General Theory of Crime.* Stanford, Calif.: Stanford University Press.

Greenberg, Stephanie, and W Rohe 1986. "Informal Social Control and Crime Prevention in Modern Urban Neighborhoods " Pp. 79–118 in *Urban Neighborhoods: Research and Policy,* edited by Ralph B Taylor New York: Praeger.

Hagan, John, and Bill McCarthy 1997. *Mean Streets: Youth Crime and Homelessness.* New York· Cambridge University Press.

Hagan, John, and Ruth Peterson, eds. 1995. *Crime and Inequality.* Stanford, Calif.· Stanford University Press.

Hanushek, Eric, and J. Jackson. 1977 *Statistical Methods for Social Scientists.* New York: Academic

Hearn, Frank 1997. *Moral Order and Social Disorder: The American Search for Civil Society.* New York: Aldine de Gruyter.

Hunter, Albert 1985. "Private, Parochial and Public Social Orders: The Problem of Crime and Incivility in Urban Communities " Pp. 230–42 in *The Challenge of Social Control: Citizenship and Institution Building in Modern Society,* edited by Gerald Suttles and Mayer Zald Norwood, N.J · Ablex Publishing.

Jacobs, Jane. 1961. *The Death and Life of Great American Cities.* New York: Random House

Janowitz, Morris 1975 "Sociological Theory and Social Control." *American Journal of Sociology* 81:82–108.

Joreskog, Karl G., and Dag Sorbom 1993. *LISREL 8: Structural Equation Modeling.* Chicago. Scientific Software International.

Kasarda, John, and Morris Janowitz 1974. "Community Attachment in Mass Society." *American Sociological Review* 39.328–39.

Kelling, George. 1998. "Crime Control, the Police, and Culture Wars: Broken Windows and Cultural Pluralism." Paper presented at the National Institute of Justice, Washington, D.C., December.

Kelling, George, and Catherine Coles. 1996. *Fixing Broken Windows: Restoring Order and Reducing Crime in Our Communities.* New York: Free Press.

King, Gary, Robert Keohane, and Sidney Verba. 1993. *Designing Social Inquiry. Scientific Inference in Qualitative Research.* Princeton, N.J.· Princeton University Press.

Kornhauser, Ruth Rosner 1978. *Social Sources of Delinquency: An Appraisal of Analytic Models.* Chicago· University of Chicago Press

Land, Kenneth, Patricia McCall, and Lawrence Cohen. 1990. "Structural Covariates of Homicide Rates: Are There Any Invariances across Time and Space?" *American Journal of Sociology* 95:922–63.

Liska, Allen, and Paul Bellair. 1995. "Racial Composition and Crime: Convergence over Time." *American Journal of Sociology* 101:578–610.

Liska, Allen, and Barbara Warner 1991. "Functions of Crime: A Paradoxical Process." *American Journal of Sociology* 96:1441–63.

Lofland, Lyn 1973. *A World of Strangers: Order and Action in Urban Public Space.* New York: Basic Books.

Mazerolle, Lorraine Green, Colleen Kadleck, and Jan Roehl. 1998. "Controlling Drug and Disorder Problems: The Role of Place Managers." *Criminology* 36:371–404.

Morenoff, Jeffrey. 1999. "Unraveling Paradoxes in Public Health: Race and Ethnic Group Differences in Pregnancy Outcomes and the Neighborhood Context of Health." Ph.D. dissertation. University of Chicago, Department of Sociology.

Morenoff, Jeffrey, and Robert J. Sampson. 1997. "Violent Crime and the Spatial Dynamics of Neighborhood Transition: Chicago, 1970–1990." *Social Forces* 76:31–64.

NORC (National Opinion Research Center). 1995. PHDCN Project 4709. Systematic Social Observation Coding Manual, June 1995. Chicago: NORC.

Park, Robert E., and Ernest Burgess. 1921. *Introduction to the Science of Sociology.* Chicago: University of Chicago Press.

Pattillo, Mary E 1998. "Sweet Mothers and Gangbangers: Managing Crime in a Black Middle-Class Neighborhood." *Social Forces* 76:747–74.

Perkins, Douglas, John Meeks, and Ralph B. Taylor. 1992. "The Physical Environment of Street Blocks and Resident Perceptions of Crime and Disorder: Implications for Theory and Measurement." *Journal of Environmental Psychology* 12:21–34.

Perkins, Douglas, and Ralph B. Taylor. 1996. "Ecological Assessments of Community Disorder: Their Relationship to Fear of Crime and Theoretical Implications." *American Journal of Community Psychology* 24:63–107

Perkins, Douglas, A. Wandersman, R. Rich, and Ralph B. Taylor. 1993. "The Physical Environment of Street Crime: Defensible Space, Territoriality and Incivilities." *Journal of Environmental Psychology* 13·29–49.

Raudenbush Stephen W 1995. "Hierarchical Linear Models to Study the Effects of Social Context on Development." Pp. 165–201 in *The Analysis of Change,* edited by J. Gottman. Hillsdale, N.J · Lawrence Erlbaum.

Raudenbush, Stephen, and Robert J. Sampson. 1999a. "Assessing Direct and Indirect Effects in Multilevel Designs with Latent Variables." *Sociological Methods and Research.* In press.

——. 1999b. " 'Ecometrics': Toward a Science of Assessing Ecological Settings, with Application to the Systematic Social Observation of Neighborhoods." *Sociological Methodology* 29.1–41.

Reiss, Albert J., Jr. 1971. "Systematic Observations of Natural Social Phenomena." Pp.

3–33 in *Sociological Methodology*, vol. 3, edited by Herbert Costner. San Francisco· Jossey-Bass

Rountree, Pamela, and Kenneth Land. 1996. "Perceived Risk versus Fear of Crime Empirical Evidence of Conceptually Distinct Reactions in Survey Data." *Social Forces* 74:1353–76.

Sampson, Robert J. 1988. "Local Friendship Ties and Community Attachment in Mass Society: A Multi-Level Systemic Model." *American Sociological Review* 53: 766–79

Sampson, Robert J., and Dawn Jeglum-Bartusch. 1998 "Legal Cynicism and Subcultural (?) Tolerance of Deviance: The Neighborhood Context of Racial Differences." *Law and Society Review* 32:777–804.

Sampson, Robert J , Stephen Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime A Multilevel Study of Collective Efficacy." *Science* 277:918–24.

Sewell, William, Jr. 1992 "A Theory of Structure: Duality, Agency, and Transformation." *American Journal of Sociology* 98:1–29.

Shaw, Clifford R., and Henry D. McKay. 1942. *Juvenile Delinquency and Urban Areas.* Chicago: University of Chicago Press.

Short, James F., Jr., and Ivan Nye. 1957. "Reported Behavior as a Criterion of Deviant Behavior " *Social Problems* 5:207–13.

Singer, Simon. 1981 "Homogeneous Victim-Offender Populations: A Review and Some Implications." *Journal of Criminal Law and Criminology* 72:779–88.

Skogan, Wesley. 1990. *Disorder and Decline: Crime and the Spiral of Decay in American Cities.* Berkeley: University of California Press

Skogan, Wesley, and Susan Hartnett. 1998. *Community Policing, Chicago Style.* New York: Oxford University Press.

Skogan, Wesley, and Michael Maxfield. 1981. *Coping with Crime: Individual and Neighborhood Reactions.* Beverly Hills, Calif.: Sage Publications.

Spergel, Irving. 1995 *The Youth Gang Problem: A Community Approach.* New York. Oxford University Press.

Taub, Richard, D. Garth Taylor, and Jan Dunham 1984. *Paths of Neighborhood Change: Race and Crime in Urban America.* Chicago: University of Chicago Press

Taylor, Ralph B. 1995. "Street Blocks with More Nonresidential Land Use Have More Physical Deterioration: Evidence from Baltimore and Philadelphia." *Urban Affairs Review* 31 20–36.

——— 1997. "Social Order and Disorder of Street Blocks and Neighborhoods: Ecology, Microecology, and the Systemic Model of Social Disorganization." *Journal of Research in Crime and Delinquency* 34 113–55.

———. 1999. "The Incivilities Thesis: Theory, Measurement, and Policy." In *Measuring What Matters* Washington, D.C.: National Institute of Justice, in press.

Taylor, Ralph B., Stephen Gottfredson, and Sidney Brower, 1984 "Block Crime and Fear: Defensible Space, Local Social Ties, and Territorial Functioning." *Journal of Research in Crime and Delinquency* 21·303–31.

Taylor, Ralph, Sally Shumaker, and Stephen Gottfredson. 1985. "Neighborhood-Level Links between Physical Features and Local Sentiments: Deterioration, Fear of Crime, and Confidence " *Journal of Architectural Planning and Research* 21:261–75.

Tilly, Charles. 1973. "Do Communities Act?" *Sociological Inquiry* 43:209–40.

Whittemore, Alice S. 1989. "Errors-in-Variables Regression Using Stein Estimates." *American Statistician* 43:226–28.

Whyte, William F 1943. *Street Corner Society: The Social Structure of an Italian Slum.* Chicago: University of Chicago Press.

Whyte, William H. 1988 *City: Rediscovering the Center.* New York: Doubleday.

Wikström, Per-Olof. 1991. *Urban Crime, Criminals, and Victims*. New York: Springer-Verlag.

Wilson, James Q., and George Kelling. 1982. "The Police and Neighborhood Safety: Broken Windows." *Atlantic* 127:29–38.

Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

Wolfgang, Marvin. 1958 *Patterns in Criminal Violence*. Philadelphia: University of Pennsylvania.

Wright, Richard, and Scott Decker. 1997. *Armed Robbers in Action: Stickups and Street Culture*. Boston: Northeastern University Press.

# Challenging the Liberal Nation-State? Postnationalism, Multiculturalism, and the Collective Claims Making of Migrants and Ethnic Minorities in Britain and Germany[1]

Ruud Koopmans
*Wissenschaftszentrum Berlin für Sozialforschung (WZB)*

Paul Statham
*University of Leeds*

As important aspects of purported tendencies toward globalization and pluralization, recent immigration waves and the resulting presence of culturally different ethnic minorities are often seen as fundamentally challenging liberal nation-states and traditional models of citizenship. According to this perspective, migrants and ethnic minorities contribute through their claims making both to the external erosion of sovereignty (the postnational challenge), and to the internal cultural differentiation of liberal nation-states (the multicultural challenge). In contrast, alternative theoretical approaches have emphasized the continuing relevance of the nation-state in the processes of inclusion and exclusion of minorities. From these three perspectives on citizenship (postnational, multicultural, and national) a set of hypotheses is derived and tested with data on the collective claims making of migrants and ethnic minorities in two European countries, Britain and Germany, for the period 1990–95. The data show very little support for the postnational approach, mixed results regarding the multicultural model, and strong support for the continuing relevance of national models of citizenship. Counter to claims that national modes of migrant incorporation have become insignificant, the evidence shows that migrant claims making is still forged in the image of a particular nation-state.

## INTRODUCTION

If we are to believe the current *communis opinio* in political philosophy and sociology, the liberal nation-state as we know it is rapidly becoming endangered. While not yet extinct, it is at least in decline or seriously

challenged. Two broad processes have been identified as driving this development (Habermas 1996). First, the nation-state's position as the predominant unit of social organization is being eroded from the outside by the forces of globalization and the shift of the locus of power from the national to the supra- and transnational levels. Second, the nation-state's legitimacy, authority, and integrative capacities are being weakened from within by the increasing pluralization of modern societies. Moreover, the liberal, universalist values that support it are being challenged by claims for special group rights (or exemptions from duties) by a multitude of collective actors who emphasize their cultural difference from the rest of society. Although the normative evaluation of these—real or supposed—trends differs widely, immigration is invariably seen as one of the main driving forces behind both the external erosion of sovereignty and the internal cultural differentiation of liberal nation-states. For Western European societies, the presence of growing numbers of racially and culturally different migrants is perhaps the most concrete, tangible, and for some most provocative, way in which globalization and pluralization have become manifest features of modern life. This holds also for the traditional immigration countries such as the United States or Australia, which have only recently become the target of larger flows of non-European immigration.

Among scholars contributing to the contemporary debates on citizenship in the migration and ethnic relations field, collective claims making by migrants and ethnic minorities are viewed as important social forces that actively contribute to and shape these trends of globalization and pluralization. The collective mobilization of political demands by migrants is often seen as a driving force that is self-evidently carrying such general transformations forward. In this article, we do not intend to enter into the extensive debates on the normative and philosophical implications of such developments, nor, obviously, do we want to discuss plural-

ization and globalization in all their facets. Our specific interest lies with
the way in which claims making of migrants and ethnic minorities of mi-
grant origin affects and is shaped by the tendencies toward pluralization
and globalization that supposedly challenge the liberal nation-state.[2] As
the theoretical debate on this topic demonstrates, citizenship—the set of
rights, duties, roles, and identities linking citizens to the nation-state—
undoubtedly plays a crucial role.[3] Three theoretical positions can be iden-
tified, two of which are directly linked to globalization and pluralization.
First, some authors (Soysal 1994; Jacobson 1996) have argued that the
postwar period has seen the rise of new forms of "postnational" citizenship
that have rendered national citizenship increasingly unimportant and are
based on the transnationalization of migrant communities and the grow-
ing role of supranational organizations and conventions that have
strengthened the rights of migrants. A second branch of scholarship sees
the increased cultural differentiation of nation-states that results from im-
migration leading to the development of a "multicultural" citizenship,
which gives special rights, recognition, and protection to minority groups
and their cultures. How far such special group rights can be reconciled
with the liberal basic values of Western nation-states is a highly conten-
tious issue (Kymlicka 1995). And third, parallel to these highly resonant
perspectives—which emphasize new forms of citizenship that undermine
traditional models of national citizenship—there is also a large number
of studies that reaffirm the importance of national models of citizenship
and that point out important cross-national differences in the modes of

[2] This implies that we limit ourselves to ethnocultural minorities of migrant origin
and exclude the problematic of so-called national minorities such as the Basques, the
Quebecois, or Native Americans. As Kymlicka (1995) has convincingly argued, such
minorities pose a rather different, and in many ways much more serious, challenge
to the liberal nation-state For similar reasons, we think our results cannot be extrapo-
lated to the situation of blacks in the United States. As descendants of involuntary
migrants and forced laborers who have suffered centuries of systematic discrimina-
tion, the position of African-Americans is fundamentally different from the (mostly)
voluntary migrants we deal with here

[3] It is important to note here that citizenship in this sense refers to much more than
just the rules for acquisition of nationality Speaking about citizenship in this narrow
sense (denoted in German as *Staatsangehörigkeit*) would not even make sense in the
British case for the period up to 1981, since strictly speaking there was no such thing
as a British nationality before the implementation of the British Nationality Act
Apart from this formal, legal dimension, our notion of citizenship (similar to the conno-
tation of *Staatsbürgerschaft* in German) includes embedded notions of the duties and
rights of citizens that are not necessarily laid down in formal law (e.g., France's strong
pressures toward cultural assimilation), as well as the more concrete policies of incor-
poration of immigrants that derive from such formal and informal notions of citizen-
ship (e.g , Britain's facilitation of the participation of ethnic minority organizations
in the political process in contrast to the marginal role reserved for foreigners' organi-
zations in Germany).

inclusion and exclusion of migrants and ethnic minorities (Brubaker 1992; Castles and Miller 1993).

Our aim in this article is to confront these three models of citizenship with data on the claims making of migrants and ethnic minorities in Britain and Germany, two of the most important immigration countries in Europe, for the period 1990–95. Both have large populations of non-European migrants and significant Muslim communities, which makes them relevant cases from the multicultural perspective. As European Union (EU) members, both countries are exposed to the process of European integration, which is invariably referred to as a paradigmatic example in studies of globalization and postnationalism. However, Britain and Germany have very divergent traditions of national identity and citizenship, which also makes them interesting for comparing the impacts of different national models of citizenship.

In the following theoretical sections we first present and discuss the three theoretical approaches in more detail. After a brief discussion of the main features of the politics of migration and ethnic relations in Britain and Germany, we derive hypotheses from the three perspectives on citizenship, regarding the claims making of migrants and ethnic minorities. We then confront these hypotheses with our data in order to assess the relative merits of the three models and to answer our central question: Does the claims making of migrant minorities constitute a challenge to the liberal nation-state, and if so, what is the nature of the challenge which it presents?

## THE POSTNATIONAL CHALLENGE

According to theorists of postnational citizenship, "Transnational migration is steadily eroding the traditional basis of nation-state membership, namely citizenship. As rights have come to be predicated on residency, not citizen status, the distinction between 'citizen' and 'alien' has eroded" (Jacobson 1996, pp. 8–9). The primary case in point are the guest workers in Western Europe. Although originally recruited on a temporary basis during the period of labor shortages between the 1950s and early 1970s, the population of guest workers and their descendants in Western Europe has continued to grow even after recruitment was formally terminated in the aftermath of the 1973 oil crisis (e.g., Baldwin-Edwards and Schain 1994). Contrary to the original rotation model, many guest workers became permanent residents and brought over their spouses and children. However, most of them did not become citizens of their new countries of residence. One of the reasons for these low naturalization rates is certainly the fact that these "denizens" (Hammar 1985) enjoy many of the same civil and social—though not political—rights as full citizens (Bauböck

1994). Postnationalists argue that these rights given to noncitizens derive from the fact that migrants have been able to sustain claims to residence and welfare rights with reference to universal rights of "personhood" based in international human rights institutions and conventions. It is argued further that these have largely taken over the role of national citizenship as the main source of rights. This has led to a decoupling of identity and rights, the two main elements of citizenship: "Rights increasingly assume universality, legal uniformity, and abstractness, and are defined at the global level. Identities, in contrast, still express particularity and are conceived of as territorially bounded. As an identity, national citizenship . . . still prevails. But in terms of its translation into rights and privileges, it is no longer a significant construction" (Soysal 1998, p. 208).

Postnationalists argue that the collective action of migrants plays an active role in the erosion and transcending of the frontiers of nation-states. The increasing speed and density of international communication and transport that are part of globalization have allowed migrants to maintain strong ties with their homelands. Migrant communities increasingly take on the character of transnationally linked diasporas and are well equipped for taking advantage of the new opportunities of postnational citizenship (Lie 1995; Shain and Sherman 1998; Jacobson 1996). The conclusion Soysal (1998, pp. 210–11) draws from these purported developments leaves no room for doubt regarding the fate of the nation-state, "In a world within which rights, and identities as rights, derive their legitimacy from discourses of universalistic personhood, the limits of nationness, or of national citizenship, for that matter, become inventively irrelevant."

Compelling as this view on postnationalism and the role of migrant minorities may be, it has not gone unchallenged. To begin with, a number of authors have pointed out that the transfer of authority to the supra- and transnational levels has not yet progressed to a level where it can seriously challenge the nation-state's prerogatives. More specifically, in migration and ethnic relations politics, the development of common EU policies and coordination of the national policies of members states has been highly restricted by the endurance of the different national policy frameworks and attachments to the "public philosophies" that underpin them (Favell 1998, p. 245; see also Joppke 1997; Freeman 1998). Even those common policies that have been implemented hardly rank as the benign human rights type that postnationalists cherish. For example, the Schengen Accord, implemented in 1995, is oriented toward better-coordinated and stricter controls on unwanted immigration.[4]

---

[4] Of course, with the 1992 Maastricht Treaty, the EU has implemented a form of European citizenship, which although still rudimentary does include complete freedom of movement and residence within the European Union for citizens of member

In addition, the practical implications of international human and civil rights conventions on the rights of migrants have also been questioned. Thus, Goodwin-Gill, Jenny and Perruchoud—representatives, respectively, of the United Nations Human Rights Commission and the Intergovernmental Committee for Migration in Geneva—conclude with regard to these treaties and conventions, "Their generality accommodates many shades of opinion, and what really counts is how the scheme of protection is worked out at the local level, particularly with regard to subsidiary rights and procedural guarantees. . . . Even under the European Convention on Human Rights the jurisprudence adopted . . . and the interpretation of 'civil rights' has sanctioned serious limitations upon nonnationals' entitlement to substantive and procedural due process. Deportation, termination of a residence permit, and the grant or refusal of entry, have all been found by the European Commission not to involve civil rights" (1985, pp. 566–68; see also Forbes 1995, pp. 198, 207).

Existing empirical evidence is rather inconclusive regarding which of these contrasting views on the role of supranational and national contexts for the claims making of migrants and ethnic minorities is closer to the truth. Against the evidence brought forward for the continuing relevance of the nation-state as a frame of reference for migrant mobilization, postnationalists point to cases where migrants frame their claims in terms of universal human rights and have successfully appealed to supranational courts to prevent or overturn national incursions on their rights (see Soysal [1994, 1997] for examples). However, such evidence remains unsystematic and often does not go much beyond the discussion of a few, supposedly representative, examples.

THE MULTICULTURAL CHALLENGE

Theories on postnationalism are linked in two important ways to multicultural citizenship. First, postnationalists argue that international conventions providing for a right to one's own culture have greatly improved the opportunities for migrants and ethnic minorities to push for the recognition of their cultural difference by the nation-state. Second, if the postnational observation is correct, that ethnic communities increasingly take the form of transnational diasporas with strong ties to the homeland, then it follows that traditional models for integrating migrants through assimilation into the majority culture will no longer work (Kymlicka 1995, p. 9).

states. However, access to European citizenship is completely dependent on access to citizenship of one of the member states. It therefore does not infringe on the prerogatives of member states to define their own rules of citizenship acquisition, nor does it alter the position of residents holding the citizenship of a nonmember state.

In the discussion of multiculturalism as a challenge to the liberal nation-state, it is not the nation-state's external sovereignty that is at stake, but its capacity to maintain social cohesion and the liberal conception of individual rights on which it rests. Again, the claims making of ethnic minorities is seen as a central aspect of this challenge: "Minorities and majorities increasingly clash over such issues as language rights, regional autonomy, political representation, education curriculum, land claims, immigration and naturalization policy, even national symbols, such as the choice of national anthem or public holidays. Finding morally defensible and politically viable answers to these issues is the greatest challenge facing democracies today" (Kymlicka 1995, p. 1). These problems are seen as arising from the fact that migrant minorities increasingly put forward demands for special group rights and recognition and support for their cultural difference and identity by the state. Such claims are often seen as challenges to the very essence of liberal values, for instance, when Muslims demand the right to polygamy, the circumcision of females, or the banning of Salman Rushdie's *Satanic Verses* (Parekh 1996). In making such demands, migrant minorities challenge the conception of a unified, undifferentiated citizenship, a development that is viewed by supporters as a healthy and necessary antidote against the prevalent "white" cultural hegemony (e.g., Young 1998) or as a serious assault on the shared communal values and solidarity necessary for social cohesion and integration (e.g., Schlesinger 1992).

We shall not dwell on the normative arguments made by these and many other participants in the debate, either for or against the granting of differential, multicultural citizenship rights to migrant minorities. Such questions are beyond the scope of this article. However, it seems to be an issue of consensus that—normatively justifiable or not, disquieting or not—claims for cultural group rights are central to the claims making of present-day migrants and ethnic minorities. For opponents of further immigration, this is a result of the (too) strong cultural differences of recent, non-European migrants, which put strain on both the migrants' adaptive capabilities and the host society's integrative capacities. From the point of view of advocates of a multicultural citizenship, it reflects the fact that exclusion in modern societies is primarily a result of discrimination or biases against groups' cultural difference, and is no longer primarily a function of socioeconomic inequality and a lack of social and political citizenship rights.

Systematic empirical evidence that would allow us to ascertain the relative importance and nature of claims for cultural group rights—both cross-nationally and compared to other types of claims—is once more in scarce supply. Unfortunately, the high level of philosophical sophistication in the debate on multiculturalism is not matched at the empirical

level. The claim that immigration and ethnic minority mobilization have led to pressures toward a differentiation of citizenship that is serious enough to potentially destabilize the liberal nation-state remains at present no more than an assertion. The usual references to ethnic strife in Bosnia, Rwanda, or the former Soviet Union are quite irrelevant to the context of immigration in Western societies. Examples drawn from more relevant contexts are usually anecdotal and extremely repetitive between studies. Open a random book on ethnic minorities and multiculturalism and you have a fair chance of stumbling across the $n$th account of the French headscarf (*foulard*) and the British Rushdie affairs. To what extent such spectacular examples are representative for the wider claims making of migrants and ethnic minorities remains unclear.[5]

## NATIONAL CITIZENSHIP REGIMES FOR INCORPORATING MIGRANTS

Recently, comparative research has refocused academic attention on the importance of citizenship rights for explaining the different national approaches in liberal democratic states for regulating immigration and the presence of foreign migrants. Not surprisingly with increasing European integration, these debates often see contemporary Western Europe as the empirical testing ground for hypotheses on why countries with ostensibly similar flows of immigration and numbers of immigrants continue to maintain different ways of attributing citizenship rights to nonnationals. An important stimulus to this lively debate on citizenship regimes and the immigration challenge to the nation-state is the work of Rogers Brubaker (1992).[6]

At a time when the rationale of EU integration calls for convergence, Brubaker points out that national regimes for incorporating nonnationals within a system of political and social rights remain, paradoxically, divergent. To explain the endurance of this national variance in citizenship regimes he draws attention to the cultural foundations of nation-states as ties that restrain and bind a community through nationhood. Brubaker's comparison of France and Germany shows that a state's legal definition for citizenship bears a distinct embedded cultural imprint of nationhood.

---

[5] Moreover, even these examples are not always well chosen. In the French *foulard* case, the overwhelming majority of public debate and attention was not created by the three schoolgirls who wore their headscarves in school, or by the Muslim organizations supporting them, but by French political leaders and intellectuals disagreeing about the nature of French republicanism and laicism.

[6] There has been a growing body of important comparative research in this field. Other examples include Kleger and D'Amato (1995), Safran (1997), Castles and Miller (1993), Cornelius, Martin and Hollifield (1994), Ireland (1994), and Kastoryano (1996).

He contrasts the *jus sanguinis* legal tradition for citizenship, where membership is acquired on the basis of ethnic descent, with *jus soli* where it is a territorial birth right. Whereas *jus sanguinis* is resistant to converts by enforcing social closure on the basis of the prescribed collective identity of national ethnicity, *jus soli* citizenship permits a contractual assimilation by an individual to nationhood, either by ritual conversion or automatically by birth. Brubaker's historical study defines Germany as a case of pure *jus sanguinis,* where citizenship rights are still based on ethnocultural belonging to nationhood. He sees France, on the other hand, as a country where citizenship is a mixture of *jus sanguinis* and *jus soli,* but where *jus soli* is strongly dominant and encoded in the Jacobin republican civic concept of nationhood.

Brubaker argues that these deeply embedded national self-definitions of citizenship continue to shape the divergent political responses of nation-states to migrants up to the present day. He cites the persistently higher naturalization rates in France compared to Germany as empirical evidence that citizenship regimes are the best explanation for the national strategies that the two countries maintain for incorporating migrants.

Extending Brubaker's analyses, a number of scholars have analyzed the citizenship configurations of nation-states as the explanatory variable for responses to the challenge of incorporating ethnic difference.[7] A considerable degree of consensus exists on the "types" of different citizenship regimes. Typically, these distinguish between two important dimensions of citizenship that can be seen to determine the degree and form of inclusiveness/exclusiveness of a national regime in relation to ethnic difference: first, the criteria for formal access to citizenship; and second, the cultural obligations that this access to citizenship entails. The first dimension relates to the distinction between an ethnocultural (*jus sanguinis*) and a civic territorial (*jus soli*) basis of criteria for attributing full citizenship, with the ethnocultural being the more "closed" and civic territorial the more "open" version. The second dimension relates to the distinction between assimilationism and cultural pluralism as the condition which a state places on attributing citizenship, with assimilationism being the more "demanding" and cultural pluralism the more "accepting." By combining these dimensions one arrives at three ideal-type citizenship regimes that can be related to specific examples of European countries: the ethnocultural exclusionist, civic assimilationist, and the multicultural pluralist.

---

[7] Castles (1995), Smith and Blanc (1994), and Favell (1998) are examples of scholars who apply national configurations of citizenship to the incorporation of migrants; whereas Koopmans and Kriesi (1997) and Koopmans and Statham (1999b, 1999c) apply a similar model by using national citizenship models as "opportunity structures" for explaining the potential for mobilization over ethnic difference—extreme right and antiracist—in different European countries.

Within the Western European context, the ethnocultural exclusionist citizenship regime is best approached by Germany. Due to the absence of access to citizenship by territorial birth and relatively high hurdles to naturalization, foreign migrants find it difficult to obtain full citizenship rights and thereby join the national community.[8] In Germany, hundreds of thousands of German-born descendants of migrants are still officially "foreigners" (*Ausländer*) without full political rights. Ethnic German immigrants (*Aussiedler*) newly arriving from Russia, on the other hand, receive full social and political rights on the basis of a hereditary link to the nation. By contrast, France has a civic-assimilationist regime, where the state offers easy access to full rights to migrants, but at the same time is loathe to accept cultural pluralism. In return for rights, migrants are expected to place allegiance to French republican values and political culture above allegiances to religious identity or ethnic belonging. French republicanism is a universalizing secular religion and ethnic collective organization has been officially accepted by the state only since 1981, which places restrictions on the form and strategic orientation of ethnic mobilization. Finally, multicultural pluralist citizenship is represented in the European context by Britain, Sweden, and the Netherlands. Here the state not only offers easy access to full social and political rights, but actually sponsors ethnic difference by recognizing immigrant groups as "ethnic minorities" with their own cultural rights and privileges. For example, in the Netherlands, recognized ethnic communities have had the right to set up their own schools and to receive funding from the state.

The key question, of course, is whether these different national traditions have significant consequences for migrants. Brubaker's answer to this question is lucidly put. According to his formula, citizenship is less 'a system for attributing rights, and more a contested political field for redefining the symbolic boundary markers for a national identity, "The politics of citizenship today is first and foremost politics of nationhood. . . . It pivots more on self-understanding than on self-interest. The 'interests' informing the politics of citizenship are 'ideal' rather than material. The

---

[8] Access to German citizenship has been made somewhat easier by legislative changes in 1990 and 1993. However, even after these changes, Germany continues to have one of the lowest naturalization rates in Europe (see Çinar 1994; Lederer 1997). Recently, however, a proposal by the new government of Social Democrats and Greens has been passed in parliament, which further lowers the barriers to naturalization and introduces a conditional form of *jus soli*, the condition being that children of migrant origin born in Germany have to give up their parent's nationality when they reach majority age in order to keep their German citizenship. The new policy, if implemented as planned from January 1, 2000 (the right has announced that it will appeal to the Constitutional Court to seek the law's repeal), would mean a clear departure from the ethnic citizenship model.

central question is not 'who gets what?' but 'who is what?' " (1994, p. 182). Here Brubaker correctly points out the importance of the significa- tion process in structuring the relationship of migrants to the nation-state. This is well demonstrated by the different labels which national policies have applied to migrants: *Ausländer* (foreigners) in Germany; *immigrés* (immigrants) in France; *etnische minderheden* (ethnic minorities) in the Netherlands. However, Brubaker's emphasis on citizenship as primarily a politics of symbolic interaction obscures the real, material consequences that such ascribed identities may have. The different symbolic labels that nations attribute to migrants directly influence the distribution of material resources to them, and their potential for mobilizing challenges and partic- ipating within the political community of a society. It makes a difference whether one is a "foreigner," an "immigrant," or an "ethnic minority"; these are not just symbols but forms of social relationships that legitimate and facilitate certain types of participation in society, while delegitimating and negatively conditioning others. Similarly, racism and discrimination is not simply the drawing of a symbolic boundary marker, but it is an action that structures the life chances of those affected, leading to system- atic (and material!) social disadvantage. Incorporating antiracist and anti- discrimination legislation into citizenship rights, and enforcing them, does affect "who gets what?" perhaps more centrally than "who is what?" Con- trary to Brubaker, there are important material interests at stake in citi- zenship rights. The resources of access to the political community—not least of which are voting rights—may indeed make a significant difference in the potential of migrant actors to mobilize and press their claims for social and political change.

However, surprisingly little systematic and comparative empirical work has been done to show how national political contexts impinge upon the collective actions and claims making of migrants and ethnic minorities (Rex and Drury 1994). One of the very few comparative studies to address empirically the important question of political mobilization is Ireland's (1994) study of forms of political activism by migrants in France and Swit- zerland (but see also contributions to Koopmans and Statham [2000]). Ire- land uses the concept "institutional channeling" for how different national forms of participatory mechanisms—in which citizenship rights are fun- damental—shape migrant political activism over time. He finds that simi- lar migrant groups mobilize differently in the two countries, which makes institutional channeling a better explanation for mobilization forms than either ethnic origin or social class.

The perspectives that we have discussed in this section claim that differ- ent configurations of citizenship are embedded in national political and civic institutional frameworks, and that these have important conse- quences both for the incorporation of migrants and for political mobiliza-

tion over ethnic difference. Before testing the relevance of these claims, it is necessary to give more contextual details on the two countries on which our empirical analysis focuses.

## COMPARING BRITAIN AND GERMANY· SIMILARITIES AND DIFFERENCES

Britain and Germany are well suited for testing hypotheses regarding the relevance of citizenship for the political claims making of migrants and ethnic minorities. Taking the three different national citizenship regimes that we have discussed, they are opposite cases. Germany is a country with ethnocultural exclusionist citizenship, where nonethnic German migrants are incorporated into the labor market, but where it is difficult for them to attain access to the political community. In contrast, Britain's citizenship model tends more toward the multicultural pluralist ideal-type, where most resident minorities have full equal social and political rights, while retaining much of their cultural difference from the majority society. To put at least some flesh on the bones of these schematic characterizations, we now look briefly at some contemporary aspects of the respective citizenship regimes, patterns and forms of immigration, as well as different policy approaches for integration.[9]

Historically, the countries have experienced broadly comparable patterns of immigration flows. Both sponsored large-scale foreign immigration due to labor shortages from the 1950s until the economic crisis in the early 1970s, when restrictive immigration controls were implemented and maintained. Many of the migrants in Germany came as a result of the "guest worker" system for importing foreign labor. This produced an official policy that assumed that these "guests" would one day return to their homelands, and so policies for integration were minimal. In contrast, many of the migrants to Britain were already subjects of the British Commonwealth and until 1971 they automatically received equal political and social rights.[10] Since 1965 race relations policies in Britain have officially taken measures to combat racism and discrimination and promote social

---

[9] Here we offer only a few background details of the country cases to assist interpretation of the data More substantive contemporary overviews can be found in Layton-Henry (1994), Solomos (1993), Mason (1995), and Modood et al (1997) for Britain, and Bade (1994), Münz, Seifert, and Ulrich (1997), and Thränhardt (1992) for Germany.

[10] The Immigration Act of 1971 and the British Nationality Act of 1981 have made "ancestry" by territorial birth—patriality—within Britain a requirement for full citizenship, which in effect prioritizes the access of predominantly white subjects of the Old Commonwealth over the predominantly nonwhite subjects of the New Commonwealth

integration, thus acknowledging that political equality has not led to full equal treatment.

In 1995, the percentage of foreign migrants relative to total population amounted to 8.8% in Germany compared to 3.4% in Britain. The higher proportion of foreigners relative to nationals in Germany compared to Britain illustrates an important outcome of the different citizenship regimes. In Germany, most migrants are counted as foreigners, as higher restrictions are placed on naturalization there than in Britain (between 8 and 15 years' residence compared to 5 years), and dual citizenship is not a possibility. Between 1990 and 1995, the proportion of naturalizations per foreign population was higher in Britain than in Germany (between 2.0% and 3.4% per annum compared to between 0.4% and 0.7% [SO-PEMI 1997]). Moreover, the *jus sanguinis* restrictions in Germany are aptly demonstrated by the fact that in 1997 the number of "foreigners" increased, despite conditions of negative net immigration. The net gain of new foreigners were born in Germany of migrant parentage. As a result of these differences, the percentage of foreigners in Britain grossly underestimates the size of the minority population. In fact, Britain and Germany have similar sizes of minority populations of migrant origin: estimates put the overall migrant population in Britain at about 8% of the total population or about 4.5 million (Castles 1995, p. 300) compared to the 8.8% or 7 million for Germany in 1995. This, of course, makes them suitable candidates for comparing the impact of citizenship configurations on migrant mobilization.

The countries of origin of the minority populations differ widely in the two countries. Today in Germany, the main foreigner groups are from Turkey (more than one-third), from the former Yugoslavia (one-quarter), and from Italy, Greece, and Poland. In Britain, a large proportion of the "ethnic minorities" are British-born African-Caribbeans (one-quarter) and South Asians (one-half). It should be pointed out, however, that both migrant populations contain ethnic and religious groups that may be taken as functional equivalents. For example, both the large Turkish community in Germany and the Pakistanis and Bangladeshis in Britain—who constitute about one-quarter of the minority population—are predominantly Muslim.

Nowadays, immigration to both countries is basically limited to family reunions and refugees, asylum seekers, or special cases.[11] An important

[11] The exception is EU citizens, who have the right to free movement and social rights within the member countries and receive local voting rights, too. Other special cases are nationally determined, such as the people of Indian origin resident in Hong Kong who were allowed to enter Britain prior to the end of the colonial rule in 1997.

exception to this rule in Germany are the immigrants of ethnic German origin (Aussiedler) who receive automatic full citizenship on arrival despite their distinct cultural difference in many cases. In the 1990s, the collapse of Communism in the East has brought new flows of Aussiedler, totaling more than 2 million in the last 10 years.

The major difference between the two countries is in the political strategies for integrating minority populations. These incorporation regimes reflect the two different traditions of citizenship to which these countries belong. As a result of their exclusion from formal citizenship, migrants and their organizations—under the label foreigners—play a marginal role in the German political process. In Britain, by contrast, migrant organization and political participation—under the label ethnic or racial minorities—is facilitated by the state. In the latter country, a state-sponsored "race relations" industry has emerged backed by antidiscrimination legislation and the authority of the Commission for Racial Equality and local bodies, which report and advise on practices for ensuring equal treatment, particularly in the labor market.[12] It is worth emphasizing that "race" is a category that British political élites adopted when attempting to address the discrimination and disadvantage suffered by the minority population. This "racialization" of policy measures in part reflected the fear of political élites that British race riots might escalate to the crisis point of those experienced in the United States. An outcome of this is that British policies were more tailored to the integration of African-Caribbeans under the generic umbrella term "black" than the relatively later inflows of migrants from the Indian subcontinent, in particular Pakistan (Rex 1991). Race relations politics has been extended to Indians, Pakistanis, and Bangladeshis under the generic term "Asian," which implies that groups with a self-identification that is nonracial, the prime example being groups with a Muslim identity, have been served relatively less well by the institutional apparatus (Modood 1988).

Another point worth making about the British case is that the constituency-based electoral system has provided the large concentrations of minority communities in specific towns and regions with a considerable resource base of voting power for influencing individual members of parliament and has in turn led political parties to take up issues relating to ethnic minorities. At the local level of politics, minorities have also made a considerable impact in specific regions (e.g., Solomos and Back 1995).

By contrast, Germany has maintained a different policy approach in

---

[12] For details on the history and implementation of British "race relations," see Solomos (1993); for a comparison with Germany and the United States, see Joppke (1996).

which migrants are excluded from political participation. Apart from the powerless and marginal foreigners' councils (*Ausländerbeiräte*) at the local level, foreigners in Germany have no institutionalized channels of access to the political process. The German state does not provide the kind of facilitation to migrant organizations that many of their British counterparts receive. There is no institutional focus for minority claims in the form of an official minority, racial equality, or antidiscrimination politics that might legitimate migrant demands and identify responsible authorities for their implementation. The official mantra ("Germany is not an immigration country") therefore has real consequences for the opportunities for minority claims making.

DATA RETRIEVAL

To investigate the political claims making of migrants and ethnic minorities in Britain and Germany, we use data drawn from content analyses of daily newspapers in the two countries.[13] In contrast to many media content analyses, we are not primarily interested in the way in which the media frame events. On the contrary, our focus is on the news coverage of mobilization, public statements, and other forms of claims making by nonmedia actors. Taking a cue from "protest event analysis" in the study of social movements (Tarrow 1989; Olzak 1989; Rucht, Koopmans, and Neidhardt 1998), our units of analysis are not articles but individual instances of claims making. Although we focus here on claims making by migrant minorities, our larger data set includes the whole spectrum of claims-making acts related to immigration, minority integration, and xenophobia irrespective of the actors involved. This includes civil society groups such as antiracist and human rights organizations or labor unions, but also political parties and state actors, including the police, courts, legislatures, local and national governments, and supranational institutions. Instances of claims making have been included irrespective of their form and range from violent attacks on other groups, public demonstrations and legal action, to public statements.

Acts were included in the data if they involved demands, criticisms, or proposals related to the regulation or evaluation of immigration, minority integration, or xenophobia. Because of our special interest in minority claims making, we included acts by resident ethnic minorities even if they were not related to these issues—provided, of course, that they involved

[13] More details on our approach and method appear in Koopmans and Statham (1999*a*).

some political claim. This allowed us to include the claims making of migrants related to the politics of their homelands.[14] Regarding territorial criteria we included acts in Britain and Germany,[15] respectively, even if they were made by foreign actors or addressed to foreign or supranational authorities, and also acts made outside Britain or Germany, but addressed to actors in these countries.[16] Both types of acts were considered to be claims in the British or German public spheres.[17]

For the comparative analysis here, we use data drawn from every second issue (Monday, Wednesday, Friday) of the *Guardian* for Britain and the *Frankfurter Rundschau* for Germany for the six-year period 1990–95.[18] These papers were chosen because they are of a comparable, moder-

---

[14] This excludes acts of international terrorism that could not be plausibly interpreted as part of the claims making of a particular resident ethnic community, e.g., most forms of Middle Eastern terrorism. Acts by terrorist groups were included, however, if they were significantly linked to a resident ethnic community, e.g., the Kurdish PKK in Germany or the planting of a car bomb by Islamic fundamentalists outside the Israeli Embassy in Britain.

[15] "Britain" here does not refer to the whole of the United Kingdom, but to the main island; events in Northern Ireland were excluded. Moreover, because Scotland has its own press, our data include few Scottish events If we talk about Britain, therefore, we are in fact mainly implying England and Wales–which is where more than 90% of the total population and an even larger percentage of ethnic minorities in Britain live.

[16] Examples of this category include the open letter by the American Jewish Committee addressed to Chancellor Kohl expressing concern about xenophobia in Germany and the claim by a Nigerian government official that the Omibiyo Family Anti-Deportation Campaign was damaging his country's national image in Britain.

[17] One may object that our focus on national public spheres introduces a bias in our results to the detriment of trans- and supranational actors. This would be true if there was such a thing as an international or European public sphere. In the absence of supranational media or transnational public debates of any importance, the public claims making of supranational actors (or claims making addressed to supranational actors) has to be mediated through national public spheres in order to be effective. At present the modest beginnings of an international public sphere in the form of television networks such as Cable News Network, or newspapers such as the now-defunct *European*, cannot compete in terms of audience or resonance with national media, not least because of the language factor, which may well prove to be an insurmountable barrier to the development of a transnational public sphere.

[18] Data were coded from microfilm (Germany) and CD-ROM (Britain) versions of the newspapers by trained coding assistants on the basis of a standardized codebook. All articles in the home news section of the newspapers were checked for relevant acts; i.e., the search was not limited to articles containing certain key words. For the main variables in the analysis (actors, addressees, aims, etc.) open category lists were used, which allow us to retain the detail of the original reports in the analysis. In addition, hard copies of the original articles were kept to allow us to go back to the original reports if information was needed that had not been captured by the variables and categories included in the codebook (codebooks are available from the authors on

ately left-liberal political affiliation and because, compared to other na-
tional quality newspapers, they have the most encompassing coverage of
the specific issues of interest. Of course, when using newspapers as a
source one has to deal with the problem of selection (not all events that
occur receive coverage) and description bias (events may get covered in
a distorted way; for these issues see McCarthy, McPhail, and Smith
[1996]). We have tried to minimize the problem of description bias by
explicitly basing the coding only on the factual coverage of statements
and events in newspaper articles, and leaving out any comments and eval-
uations made by reporters or editors. In any case, quality newspapers have
to protect their reputation and cannot afford to quote claims patently in-
correctly. Since our interest here is in publicly visible claims making, the
problem of selection bias is less aggravating than in some other contexts,
because acts of claims making become relevant—and potentially contro-
versial—only when they reach the public sphere.

However, it may be that our sources have specific biases that make the
data drawn from them unrepresentative for the print-media landscape at
large. To check for such biases, we have in each of the two countries
drawn additional samples from other newspaper sources. In Germany,
we coded two years of the right-wing tabloid *Bild Zeitung*, one year of
the German-published edition of the Turkish daily *Hürriyet*, as well as
four-month samples from three different local dailies. In Britain we have

---

request). The use of very detailed open category systems including hundreds of differ-
ent actors and claims entails that conventional measures of intercoder reliability are
not applicable to these variables. The categorizations used in this analysis are not
based on coder decisions but are the result of aggregations of raw codes by the authors
(for a similar two-stage procedure of content analysis, see Shapiro and Markoff [1998],
chaps. 6, 11). Conventional reliability measures are, however, applicable to one impor-
tant aspect of the coding process, namely the inclusion or exclusion of articles and
claims from the newspaper source (we report results for intercoder reliability tests
based only on the German case, since the British data were collected by one single
coder and thus intercoder reliability is irrelevant here). In a first step, coders have to
decide which newspaper articles contain any codable claims at all. This is a crucial
step in the coding process because only these articles will be archived, and going back
to the original source material for checks, corrections, and refinements will not be
possible for any articles that are excluded at this stage. Comparing six coders regarding
their inclusion or exclusion of articles, a reliability coefficient (Cronbach's alpha) of
.95 was achieved. In a second step, coders have to identify claims, the final unit of
analysis, in the selected articles. Often, articles contain several codable claims, which
have to be identified and delimited from each other. This increases divergences among
coders so that in the final instance we arrived at a reliability coefficient of .92 in our
test. Unlike the lost information in the first step, coder unreliability in this second
step of the coding process is reparable. For the present analyses, e.g., we have checked
many of the original source articles in order to improve our qualitative understanding
of minority claims making. In the process, we have found some coding errors and
have been able to correct them.

a cross-section of six national newspapers for the year 1995.[19] Comparisons among these sources confirm that our two main sources give relatively broad coverage to issues of immigration and ethnic relations. It is important to note that the sometimes large differences in coverage rates do not lead to very different distributions of acts on important variables. As an example we may take the *Frankfurter Rundschau* and the *Bild Zeitung*, which occupy diametrically opposed positions in the German print-media landscape, on both the left/right, and quality/tabloid dimensions. The number of reported claims in the domain of immigration and ethnic relations turns out to be 4.6 times higher in the *Rundschau* than in the *Bild Zeitung*. However, distributions across different issues (asylum, integration, antiracism, etc.) hardly differ among the two papers and neither does the representation of different actors in the coverage.[20] In as far as minor differences exist, they run in a direction that is helpful rather than harmful for our purposes. Thus, the *Bild Zeitung* has a somewhat stronger focus on national actors to the detriment of regional and local ones and a somewhat weaker coverage of acts by minority actors (4.5% vs. 6.8% for the *Rundschau*). Our main source thus is somewhat less selective when it comes to covering the claims making of more marginal actors.

These similarities and minor differences concur with large differences in the evaluation of events, as is revealed by a comparison of editorials in both papers. Taking the issue of asylum, 70% of *Bild Zeitung* editorials favored restrictions in the rights of asylum seekers, while 70% of *Rundschau* editorials opposed such restrictions. However different these newspapers' political stances may have been, they had marginal effects on coverage. Of statements by other actors reported in the *Bild Zeitung* an average of 56% were in favor of restrictions, against 48% for the *Rundschau*. We may conclude, therefore, that our strategy of focusing on factual coverage and ignoring newspapers' framing and commenting of events reduces biases that may effect the variables we are interested in to such an extent that we can safely conclude that our main sources give a representative picture of public claims making on immigration and ethnic relations.[21]

[19] Apart from the *Guardian*, these were the *Times*, the *Daily Express*, the *Daily Mirror*, the *Sun*, and the *Daily Mail*.

[20] This becomes very clear if we compare actors with known political party affiliation. The distribution is strikingly similar in both papers, and, interestingly, closely matches the electoral strength of each of the parties.

[21] A comparison of the *Rundschau* to the *Hurriyet* for the year 1995 produces similar results. As a mainstream newspaper with a readership drawn mainly from the majority culture the *Rundschau* of course does not cover claims making by ethnic minorities to the same extent as the *Hurriyet*, which caters specifically to the Turkish immigrant community. Indeed, the number of ethnic minority claims reported was about three times higher in the *Hurriyet*, with, of course, a heavy bias toward claims making by

HYPOTHESES

In the remainder of this article, we will use our data on migrant and ethnic minority claims making to test the validity of the three theoretical views on the link between citizenship and migration. Although these perspectives are not as such theories of ethnic mobilization, minority claims making figures centrally in theories of postnational and multicultural citizenship, both as a cause and as a consequence of shifts in the nature of citizenship away from the traditional unitary form of national citizenship. While the national citizenship model discusses ethnic mobilization less centrally, it also implies clear expectations about the nature of, and cross-national differences in ethnic minority claims making. Thus, each of the three theoretical configurations of citizenship is associated with a particular pattern of minority claims making.

As we have seen, migrants and ethnic minorities are seen by those who proclaim the advance of a postnationalist or multicultural citizenship as actors who, through their claims making, make important contributions to these trends. Thus a high level of "postnational" or "multicultural" claims making by migrants and minorities may be regarded as direct supportive evidence for these models. As indicators of postnational claims making we will look at the prevalence of transnational collective identities and organizations, the degree to which claims are addressed at authorities beyond the nation-state, and the extent to which the substance of minority claims refers to supra- or transnational institutions and conventions. As indicators of multicultural claims making we regard the degree to which minorities collectively identify themselves along cultural (particularly religious) lines and the extent to which their claims making focuses on extensions of cultural rights.

The three modes of citizenship can also be seen as political opportunity structures (McAdam 1982; Tarrow 1994; Kriesi et al. 1995) that shape the mobilization of migrants and minorities by facilitating some and constraining other forms of claims making. Thus, the increased institutionalization of postnational and multicultural institutions, rights, and dis-

---

Turkish and Kurdish groups. Again, however, this difference in the quantity of coverage hardly affects the distributions among types of claims. As we will see further on, the most striking characteristic of German minority claims making is the predominance of homeland-related claims. The comparison with the *Hürriyet* shows that this result is not the product of a construction by German mainstream media, who would ignore claims making by minorities for rights in Germany and focus disproportionally on conflicts imported from migrants' homelands. For the year 1995, the percentages of homeland-related claims in the two newspapers are very close: 67% in the *Rundschau* and 60% in the *Hürriyet*. Information on similar checks for biases in the British case, or for the comparison of national and local newspapers are available from the authors on request.

courses implies increased opportunities for migrants and minorities for advancing postnational and multicultural claims of the kind indicated above. On the contrary, the national citizenship model implies that migrant and minority claims making will be strongly affected by the opportunities and constraints for intervening in the public sphere set by the institutions, rights, and discourses that derive from nationally specific modes of incorporation of migrants into the political community through citizenship. From these general expectations we can draw a number of concrete hypotheses relating to specific aspects of migrant and ethnic minority claims making:

Hypothesis 1

At the most general level, the national citizenship model leads us to expect important differences in patterns of claims making of migrant minorities between Germany and Britain as a result of the different ways in which migrants and minorities are incorporated into or excluded from the nation-state. The postnational citizenship model, by contrast, leads us to expect similarities between the two countries because of the fact that migrants' claims making increasingly depends on the opportunities provided by universal human rights, which are defined and encoded at the supranational level and thus are not fundamentally different for minorities in Britain or in Germany. Although theories of multicultural citizenship are less explicit in this regard, they also lead us to expect cross-national similarities rather than differences. The affirmation of cultural difference by ethnic minorities and the concomitant rise of claims to differential citizenship rights are seen as trends affecting modern Western societies in general.[22]

Hypothesis 2

The type of minority actors involved in claims making may be distinguished along two dimensions of theoretical relevance: (2.1) the kind of collective identities they mobilize, and (2.2) the territorial scope of their organizations.

*Hypothesis 2.1.*—The first actor dimension refers to collective identities. The kind of collective identities that migrant minorities put forward

---

[22] Significant cross-national variation could result, in this view, from differences in the composition of the migrant population, leading to different degrees of cultural difference between minorities and the majority society. However, since Britain and Germany have comparable numbers of non-European and Muslim minorities, we should not find large differences between these two countries.

in their claims making are not self-evident, but signify specific construc-
tions of the group's image of itself and its relation to the wider society,
including other migrant groups. Moreover, identities and labels may be
strategically used in such a way that the impact and legitimacy of claims
is optimized. Since the multicultural model sees cultural difference as the
primary source of minority claims making, we would expect cultural and
religious forms of identification to dominate. The national citizenship
model would again expect different outcomes in different national con-
texts, depending on which collective identities are sponsored and ex-
cluded, respectively, by the national model of migrant incorporation. In
the British context, this leads to the expectation that migrants will primar-
ily make claims on the basis of identification with one of the two officially
recognized racial groups, "blacks" and "Asians." To a lesser extent, British
multiculturalism may also further claims making on a cultural or religious
basis. Homeland-based forms of identification should be relatively rare
in Britain. In Germany, however, we would expect the collective identities
of "foreigners" to match precisely homeland-based categories, either na-
tionality (e.g., the Turkish community in Berlin),[25] homeland-based politi-
cal affiliation (e.g. the Turkish Communist Party), or homeland-based eth-
nic cleavages (e.g., Kurdish nationalist groups).

  *Hypothesis 2.2.*—Regarding the territorial scope of migrant organiza-
tions, postnationalism theories, which emphasize the transnationalization
of ethnic communities, predict a significant involvement of actors that
transcend the national context. These may be either international (e.g.,
the World Council of Roma and Sinti), Europewide (e.g., the European
Association of Turkish Acadamics), homeland-based (e.g., the National
Resistance Council of Iran)[24] or otherwise foreign-based (e.g., the Ameri-
can Jewish Committee). The national citizenship model, by contrast, does
not see an important role for transnational actors. The involvement of
homeland-based actors, however, would depend on the kind of national
citizenship regime. In countries such as Germany, where immigrants con-
tinue to be defined as foreigners and have few opportunities to participate
in German political life, we might expect migrants to remain strongly tied
to their homelands, which might in turn strengthen the position of home-
land-based organizations. In more inclusive countries such as Britain,
where most migrants are citizens and the state sponsors migrant and mi-

---

[23] All examples of organizations and claims given in this and the following section
were drawn from our data.

[24] As this example indicates, our definition of homeland-based includes exile organiza-
tions of opponents to the homeland regime.

nority organizations that operate within the British political process, we might expect homeland-based organizations to play a relatively modest role.

## Hypothesis 3

We can formulate hypotheses about the addressees of claims if we take into account the three perspectives on citizenship (postnational, multicultural, and national).[25] When looking at addressees, we can formulate hypotheses similar to those for the territorial scope of migrant organizations. Once more, the postnational model leads us to expect that to an important extent migrant minorities direct their claims at authorities that transcend national borders. These may be either international (e.g., the UNHCR), European (e.g., the European Parliament), or represent a foreign state (e.g., the Turkish government). By contrast, from the national citizenship model we may draw the expectation that political authorities within the national polity are the most important addressees of claims. In Germany we may in addition expect a significant number of claims addressed at the authorities of migrants' homelands, linked to the persistence of homeland-based organizations and collective identities discussed above.

## Hypothesis 4.1

Regarding the content of claims made by migrants and ethnic minorities, we ask a number of different questions. First, we look at the relative importance of migrant claims making on immigration and ethnic relations compared to migrant claims making on homeland issues and all claims making on immigration and ethnic relations. For the relation between migrants' claims making on issues related to the countries of residence and of origin, respectively, the national citizenship model allows us to derive a clear expectation.

We might expect the importance of homeland-related issues to depend on the degree to which migrants are symbolically included in the national political community and materially incorporated into the political process of their countries of residence. Thus, we expect a stronger emphasis on homeland issues in Germany, and a stronger orientation toward issues

---

[25] By the term "addressee" we refer to the authorities at whom minorities direct their demands, policy proposals, or criticisms. In other words, the addressee is the political actor that is called upon to act on behalf of the claims makers or to refrain from acting against their interests.

relating to integration into the host society in Britain. We shall interpret a strong and relatively similar representation of homeland issues in both countries as support for the postnational model. It is plausible to expect that the transnationalization of ethnic communities also leads them to retain a substantive interest in the politics of their homeland countries.

## Hypothesis 4.2

Regarding the contribution of migrants to the overall claims making on immigration and ethnic relations, both the postnational and multicultural model hypothesize a relatively important role for migrants and ethnic minorities. After all, their claims making is supposed to be an active force in the emergence of new, postnational, or multicultural forms of citizenship that challenge the liberal nation-state. The national citizenship model would again say that it depends on the national context. In inclusive countries such as Britain, where minorities are citizens and their involvement in the political process is facilitated by the state, migrant and minority claims making might be expected to be relatively important. Where migrants lack political rights and the state does not facilitate migrants' political organization and incorporation into the policy process, as in Germany, migrants might be expected to remain relatively marginal actors.

## Hypothesis 4.3

Finally, we look at migrants' claims on immigration and ethnic relations in more detail. We first ask to what extent these claims are framed with reference to supranational, and transnational, or national institutions, rights, and conventions. Second, we look at the type of demands and particularly at the nature and relative importance of demands for cultural group rights.

With regard to the territorial scope of claims, we can draw a clear hypothesis from the postnational model. According to this perspective, the relativization of the nation-state occurs not only because foreign, transnational, and supranational actors intervene in national politics (hypothesis 2.2) and actors in the national polity directly address supranational or foreign authorities (hypothesis 3). At least as important as these direct forms of "postnationalization" is the indirect form in which actors make claims on national authorities by referring to supranational institutions, conventions, and legislation. The postnational model predicts that such claims play an important role, while the national citizenship model would expect them to be relatively marginal compared to claims whose frame of reference remains within the national context.

Hypothesis 4.4

Finally, we look at the type of demands that minorities make with regard to immigration and ethnic relations. Here only theories of multicultural citizenship allow us to formulate clear expectations. According to this perspective, demands for special group rights related to the recognition and protection of cultural differences are central to the claims making of ethnic minorities. The national citizenship model does not allow us to say much about the relative importance of such claims compared to other types of claims. However, they may be expected to play a more important role in Britain than in Germany, because the former country officially sees itself as a "multicultural society" with tolerance for cultural diversity as one of its constituent principles. Thus, opportunities for claims making on the basis of cultural difference would seem to be more favorable in Britain than in Germany.

FINDINGS

Starting with the type of actors involved in minority claims making, table 1 gives an overview of the type of collective identities that were expressed in claims making by ethnic minorities and migrants (hypothesis 2.1). We distinguish four broad types of identities. First, migrants and minorities may identify themselves across ethnic and cultural boundaries on the basis of their common status as "immigrants," (ethnic) "minorities," or "foreigners." More specific status-group identifications include asylum seekers or the ethnic German immigrant category of Aussiedler. Second, minorities may identify, or be identified with a certain racial group, such as blacks or Asians. As a powerful cultural marker, religion can be a third possible basis for migrant and minority claims making. Finally, migrants may identify themselves on the basis of their common national or ethnic descent. It is well worth noting that these possible types of identification overlap and are to an important degree in competition with each other. Whether immigrants from Bangladesh identify themselves as immigrants, as Bangladeshis, as Muslims, as Asians, or as blacks is a political outcome that may give us important information on the nature of the relation between immigrants and the nation-state. As we have indicated in hypothesis 2.1, the multicultural model leads us to expect identification on the basis of cultural difference, whereas the national citizenship model predicts identification on the basis of national or ethnic origin in Germany and on the basis of racial and cultural groups in Britain.

If we look first at the first and third columns of table 1, we see that indeed identification with national or ethnic groups is much more frequent in Germany (83%) than in Britain (19%). It is interesting that, apart from some claims making by Jewish religious groups, we find almost no claims

TABLE 1

COLLECTIVE IDENTITIES OF ACTORS IN CLAIMS MAKING BY MIGRANTS AND ETHNIC
MINORITIES IN GERMANY AND BRITAIN, 1990–95 (%)

| | GERMANY | | BRITAIN | |
|---|---|---|---|---|
| | All Claims | Named Organizations | All Claims | Named Organizations |
| Status groups: | 14 1 | 8.0 | 19.1 | .0 |
| Foreigners ........... . . .... .. ....... . | 5.3 | 5.7 | 4 | . . . |
| Asylum seekers ... . .. .. .............. . | 7.7 | .8 | 10 2 | . |
| Immigrants ............ .. . .. ................... | . . | . . . | 4 | . . |
| Minorities ... ... . ...... ... ...... . . . .... | .3 | .5 | 8.1 | . |
| Aussiedler ... ... ... . ... ............. . . | .8 | 1.0 | . . | . . |
| Racial groups: | .1 | .0 | 37.7 | 34.2 |
| Blacks ..... .. . . ............. . .... .. ......... | .1 | . . . | 26 3 | 34.2 |
| Asians ... ... ... . ......... ...... ..... .. . ...... | . . | . . . | 9 3 | . . . |
| African-Caribbeans ........ ......... . .... | . . | . . | 1.3 | . |
| Black African ......... . .. . ............... . | . . | . | 8 | . |
| Religious groups: | 3.3 | 6.4 | 24 6 | 47 4 |
| Religious Jewish ......... .. .. ...... .. ........ | 2.6 | 5.1 | . . . | . . . |
| Muslim/Islamic ... . .... ......... . . . | .4 | .8 | 24.2 | 47.4 |
| Other . .... ... .... ... . .. ..... ..... .. . . | 3 | .5 | .4 | . . |
| National and ethnic groups | 82 5 | 85.6 | 18 6 | 18 4 |
| EU countries ... .. . ................ . . | .5 | .3 | .4 | 1.3 |
| Ex-Yugoslavs ... ... ...... . . ....... .. .... | 1.7 | 1.5 | 4 | . |
| Roma and Sinti/gypsies ........... . .. . . | 11.2 | 14.4 | .8 | 2.7 |
| Secular (ethnic) Jewish . ..................... | 17.6 | 36 5 | 4.7 | 7 9 |
| Other European ............ ... .. .. .......... | .1 | . . | .8 | . . . |
| Turks ... . . .. ......... .. .. .... .. ........ | 15 2 | 14.1 | 4 | . . . |
| Kurds ............. ... . . ... .. ............ ... . | 32.9 | 20.1 | .8 | . . . |
| Iranians .. ............. .. . .... ............. . . | 2 0 | 2.8 | 4 | . . . |
| Pakistanis . ............ ... ..... ... .. .. ..... | . . . | | 1.3 | 2.7 |
| Bangladeshis ............ . .. ............ ......... | . . . | . . | 4.2 | 2.7 |
| Indians ... .. . ............ . .. ........... ........... | . . | . . | 1.3 | 1.3 |
| Sikhs .. . .. . . ....................... ... .. .. | | | 1 3 | . . . |
| Others . . . .. ......... ...... ..... . ........... | 8 | 1.1 | 1.8 | . |
| Total .. .. . .. . ..... .. ............... . . | 100.0 | 100.0 | 100 0 | 100 0 |
| N ... . ..... ............. .. . .. .. ..................... | 757 | 389 | 236 | 76 |

NOTE.—$\chi^2 = 491$ 17, $P = 0$ 0, $df = 3$ calculated on the basis of main groupings (e g , "status groups,"
"racial groups," etc ), $\chi^2 = 259.17$, $P = 0$ 0, $df = 3$ for named organizations.

making on a religious basis in Germany.[26] For a country with a Muslim population of several million it is particularly striking that there were very few (0.4% or $N = 3$) instances of claims making by Muslim or Islamic groups. While Muslim minorities in Germany enter the public sphere as Turks, Kurds, Bosnians, or Iranians, rather than as Muslims we find the exact reverse pattern in Britain. While 24% of claims making originated in Muslim groups, only 1% and 4%, respectively, were made under the label of the two national groups, Pakistani and Bangladeshi, to which most British Muslims belong.

We find a similarly striking difference between Germany and Britain regarding the prominence of racial groups in claims making. For obvious historical reasons related to the race politics of the Nazi period, race has never gained currency in postwar German political discourse and has no place in the state's aliens politics.[27] By contrast, race is a legitimate and regularly used basis for claims making in Britain, where 26% of claims originated in black and another 9% in Asian groups.[28]

An important objection that may be raised against these findings is that because they are based on media data they are likely to measure the ascription of identities by the majority society, which, interesting as it may be, is not necessarily the same as the self-identification of migrants and ethnic minorities. We therefore also present a second measure of collective identities (displayed in table 1, cols. 2 and 4), which is based on the subsample of cases for which we had information on the full name of the organization that made the claim. This excludes cases in which the newspaper only gave a vague identification of the claims makers (e.g., "Turkish groups" demonstrating against racism), which leave considerable room for the ascription of identities by journalists. Names of organizations, by contrast,

[26] In the case of Jews and Sikhs it is often difficult to tell whether the label used refers to ethnicity or to religion. Both in English and in German the same word, Jewish (*jüdisch*) is used to denote Jews both as a religious group and as an ethnic group. We have used the religious group category only if the name of an organization indicated a religious basis (e.g., Synagogal Society Adass-Jisrael), while the ethnic category was used otherwise (e.g., the Central Council of Jews in Germany—a secular organization that was responsible for the large majority of claims by Jews in Germany). This problem may have caused a minor underestimation of the religious component in the mobilization of Jews (for Sikhs we have only three cases, so the question does not have much practical relevance).

[27] Even racism is usually not referred to in Germany as such, but as "hostility against foreigners" (*Ausländerfeindlichkeit*).

[28] Note that the label "black" has two meanings in the British context, which our data do not allow us to disentangle. Black in the narrow sense of the word, as used in the state's race relations politics, refers to people of African-Caribbean and sometimes also of African origin. However, there is also a broader, political meaning of black, which refers to (non-European) ethnic minorities in general (including Asians), and has become popular among radical minority and antiracist groups.

are important vehicles for the self-presentation of groups toward both their constituency and the wider society, and therefore may be considered good indicators of the group's collective identity. For the German case, this different method does not lead to important changes. Asylum seekers almost completely disappear and the share of Jewish groups increases, but this is simply a result of the low degree of organization of the marginal group of asylum seekers, and the strong institutionalization of the-state-sponsored-Jewish community. Regarding theoretically more important findings, however, nothing changes: national origin becomes even more important as a basis for collective identities (86%) and racial and religious identifications remain extremely marginal.

For Britain, the general pattern also remains much the same, but there are two significant, and probably interrelated, changes. First, unlike the category "black," the racial category "Asian" does not seem to be relevant for self-identification. In fact, we found no organizations at all which identified themselves as Asian. By contrast, the share of the religious category "Muslim" almost doubles when we consider only named organizations. Since Asian and Muslim are alternative identifications for largely the same immigrant groups (Pakistanis and Bangledeshis), it may be the case that the difference between the results in columns 3 and 4 results from media bias, which—in line with the official categories of British race relations politics—describes self-identifying Muslim claimants as Asians. This fits other accounts of the problems British Muslims encounter in their claims making on a race relations system that was modeled on African-Caribbean blacks and has difficulty incorporating Muslim immigrants from South Asia (see, e.g., Modood 1988).

Overall, the comparative results provide strong support for the national citizenship model. Neither the postnational nor the multicultural model is able to explain why British minorities identify as racial or religious groups and their German counterparts identify on the basis of their homeland national or ethnic origin. Taken on their own, the British results are in line with what the multicultural model would lead us to expect, but the comparison with Germany shows that the strong mobilization of religious identities in the British context is not a general consequence of the presence of culturally different minorities but depends on a facilitating political context, which stimulates claims making on the basis of cultural identities. Where such an affirmation of multiculturalism by the state is lacking and immigrants are officially seen as citizens of another state, as in Germany, national origin becomes the overriding form in which migrants are identified and identify themselves.

One finding for Britain, however, is difficult to reconcile with the national citizenship model and supports the multicultural emphasis on cultural difference as an important source of migrant claims making. Al-

TABLE 2

TERRITORIAL SCOPE OF ORGANIZATIONS IN CLAIMS MAKING
BY MIGRANTS AND ETHNIC MINORITIES IN GERMANY AND
BRITAIN, 1990–95 (%)

|  | Germany | Britain |
|---|---|---|
| Supra/transnational | 5 | 7.2 |
| European | .5 | |
| Foreign-based | .3 | 1.2 |
| Homeland-based | 26.0 | 6.0 |
| National | 72.7 | 85.5 |
| Total | 100.0 | 100.0 |
| N | 388 | 83 |

NOTE.—$\chi^2 = 33\ 55$, $P = 0\ 0$, $df = 4$

though the racial label "Asian" is the officially sponsored identity for migrants from the Indian subcontinent, the Muslim identity of Pakistani and Bangladeshi immigrants is much more powerful, especially in self-identification. Although the British variant of multiculturalism cannot be seen as a political environment hostile to the mobilization of cultural identities, British race relations politics certainly favors "black" or "Asian" over "Muslim." The dominance in migrant claims making of "Muslim" over "Asian" therefore indicates that there are limits to the capacities of national modes of minority incorporation to shape migrant identities in their own image. And likewise it shows that there is at least some truth in the multicultural emphasis on cultural difference.

A second aspect of the actor dimension concerns the territorial scope of migrant and minority organizations (hypothesis 2.2). For those acts of claims making for which an organization was explicitly mentioned, table 2 shows whether these organizations were international, Europewide, homeland-based, or otherwise based in a foreign country, or whether their territorial scope remained confined to the German and British polities. Contrary to the hypothesis drawn from the postnational model, the territorial scope of the large majority of migrant and minority organizations does not reach beyond the boundaries of the nation-state: 73% of all claims in Germany and as much as 86% in Britain were made by national organizations of immigrants and minorities, based in, and limited to the German and British polities.[29] In contrast, international, European, and foreign

[29] The label "national" here and in the remainder of this section refers to "within the confines of the nation-state" and therefore includes the national, regional, and local levels of the polity

migrant organizations played a marginal role in claims making. In the light of all the talk about Europe, both in the postnational perspective and in the wider academic discussion, the most surprising finding is perhaps the virtual absence of European-level migrant organizations. In Britain we found no claims by such organizations at all, in Germany just two cases. This finding confirms the point of view of those authors who have pointed out that migrant organizations on the European level have remained relatively impotent actors, not least because migrant groups—even from the same ethnic origins—from different European countries often have widely diverging opinions about the aims and strategies of integration and antiracism (see, e.g., Favell 1998).

Of course, for those bent on looking for evidence of postnational trends, there were a few examples of the involvement of transnational organizations in claims making. The data show, however, that such cases were the rare exception rather than the rule. The only (at least in Germany) significant form of organization that transcended national boundaries were organizations based in migrants' homelands such as the Kurdish Workers' Party (PKK) or the Iranian People's Mojahedin. To the extent that they put forward claims aiming at the improvement of the situation of migrants from the respective homeland in Germany or Britain, such forms of migrant organization may still be interpreted as evidence of postnational tendencies. However, such cases were rare, too. Homeland-based groups were almost exclusively occupied with political issues relating to the politics of migrants' countries of origin: for example, Kurdish separatism, rivalry between Turkish extreme right and left groups, or opposition of Chinese exiles against the Communist regime in China. Far from constituting some kind of challenge to the nation-state, such forms of claims making reflect the ongoing conflicts surrounding state-building and consolidation in many countries outside Western Europe and North America.

An additional weakness of the postnational model is that it has no answer to the question why homeland-based organizations play a much more important role in Germany than in Britain (26% and 6%, respectively). As we have indicated above, this result can be explained within the national citizenship model by the different ways in which migrants and minorities are included or excluded from the political community. As we saw in table 1, the exclusive German model of ethnic citizenship and the labeling of migrants and minorities by the state as foreigners has led migrants in Germany to retain a collective identity based on the national and ethnic categories of their homelands. Such identifications, of course, benefit homeland-based groups, who—like the German state—continue to see migrants as citizens of their countries of origin who can be called upon to make a contribution to the political struggle in the homeland. By contrast, the British model facilitates migrant and minority organization

TABLE 3

TERRITORIAL SCOPE OF ADDRESSEES OF CLAIMS MAKING BY
MIGRANTS AND ETHNIC MINORITIES IN GERMANY AND
BRITAIN, 1990–95 (%)

|  | Germany | Britain |
|---|---|---|
| Supra/transnational ..... . ...... ...... . ..... .... | 1.1 | 2.2 |
| European . .... ....... . ..... ..... .. ...... . ..... .... | 1.1 | 7 |
| Homeland governments ..... ..... ...... .... | 24.5 | 2.2 |
| National authorities .. ............... ..... .. .... | 73.3 | 95.0 |
| Total ..... . ....... ..... . ..... . ..... ..... .. ... | 100 0 | 100.0 |
| N .. ...... ..... ...... . ...... ....... ....... ..... .. .... | 351 | 141 |

NOTE —$\chi^2 = 34.60$, $P = 0 0$, $df = 3$

within the national context, both because it actively facilitates such forms
of organization through financial support and the opening of channels of
access to the political process and because it stimulates migrants to iden-
tify themselves as racial or cultural minorities within British society.

The postnational model does not stand or fall, however, by the transna-
tional organization of collective action alone. One form in which postna-
tional citizenship takes shape is when collective actors, whether them-
selves transnational or not, bypass national authorities and directly
address institutions and authorities outside the nation-state (hypothesis
3). Our data contain several examples of claims making along these lines.
Thus, in 1994 a group of 200 German Kurds drove to French Strasbourg
to offer a petition to the European Parliament against the persecution of
Kurds in Turkey. In another example, German Roma and Sinti groups
appealed to the UNHCR to move against the German government's plan
to deport refugees from this ethnic group back to Romania. In Britain, we
have two cases in which the International Islamic Front and the Supreme
Council of British Muslims called upon the Allied Coalition during the
1991 Persian Gulf War to cease hostilities. But there is precious little else.
While the postnational model suggests that these are typical examples of
migrant claims making in the age of globalization, table 3 makes clear
that such postnational claims making was highly exceptional. Again, the
most surprising finding may be the insignificance of the EU and its institu-
tions as a target for migrant and minority demands (about 1% of claims
in both countries). With 2% in Britain and 1% in Germany, claims ad-
dressed to other supranational institutions were equally rare. By contrast,
the nation-state was the target of 73% of German and as much as 95%
of British minority claims making. Once more, the only significant form
of claims making transcending national borders are claims addressed at

TABLE 4

MAIN ISSUE FIELDS OF CLAIMS MAKING BY MIGRANTS
AND ETHNIC MINORITIES IN GERMANY AND BRITAIN,
1990–95 (%)

|  | Germany | Britain |
|---|---|---|
| Immigration politics ..... . .... .. ..... .. ..... . . | 1.5 | 5.9 |
| Asylum politics ...... .. .. ... . ...... .. ... . ...... | 15.7 | 11 9 |
| Minority integration politics ...... .. ... ... | 6 3 | 50.8 |
| Antiracism/xenophobia . ...... .... .... .. .. | 32.1 | 21.6 |
| Interethnic conflict . ...... .. ..... . ...... . ... . .. | .9 | 3.8 |
| Homeland politics ... . ...... . ....... ..... ...... | 41.5 | 4 2 |
| Other . ...... .. ...... .. ...... .... . .... . ...... .. ... | 2 0 | 1.8 |
| Total ...... ....... ..... ..... ...... .. ... . ... | 100.0 | 100 0 |
| N .. ...... . ..... .. .. ... ... . .... .. ... ...... .... | 757 | 236 |

NOTE.—$\chi^2$ = 319 50, $P$ = 0 0, $df$ = 6.

the governments of migrants' homelands. Such cases are not of the type asking homeland governments to intervene with the German or British governments on behalf of migrant rights, which would still fit the postnational model. Almost exclusively they are related to political conflicts in the homeland. Finally, we note again that homeland-related claims are much more frequent in Germany than in Britain, which provides further support to the comparative arguments drawn from the national citizenship model discussed above.

Having discussed collective identities, organizations, and addressees, we now turn to the content of claims by migrants and minorities. Table 4, which shows the distribution of migrant and minority claims across main issue fields, adds further detail to the picture. Apart from immigration and asylum issues, which are roughly of the same importance in both countries, Britain and Germany provide a striking contrast. German minorities made claims on homeland issues 10 times more often than their British counterparts (42% compared to 4%). In Britain, most claims fall in the category of "minority integration politics," by which we denote claims making related to the integration, rights, and social position of resident minorities, which, by contrast, make up only a small minority of German claims (51% compared to 6%). The issue field of antiracism, which includes claims making against xenophobia and right-wing violence,[30] is somewhat, though not dramatically, more important in Germany (32%

[30] Claims against institutional forms of discrimination, e.g., by the police or the judiciary, were included in the category of minority integration politics.

against 22%). In line with the national citizenship model (hypothesis 4.1) we thus find that British minorities focus primarily on their integration and rights within British society, whereas the most important field of claims making for their German counterparts refers to the political situation in migrants' homelands. In as far as German minorities make claims related to their position in Germany, these tend to be largely defensive: most of the antiracist demands were protests against the wave of xenophobic violence that swept across Germany in this period, sometimes accompanied by the demand on German authorities to provide better protection against such attacks. Proactive demands for an extension of migrant and minority rights were, however, not very frequent in Germany, in stark contrast to Britain. We will return to this in more detail below.

For the moment, we would like to point to an interesting finding in the German case that emerges from a comparison of the percentage for homeland issues in table 4 (42%) and the percentage of claims addressed at homeland authorities in table 3 (25%). This difference is partly explained by claims for which no explicit addressee was mentioned. These were primarily attacks by Kurdish groups against Turkish targets in Germany, such as restaurants, banks, or travel agencies. Indirectly, such acts of violence of course also aimed at putting pressure on the Turkish government, but because this was not explicitly mentioned they do not show up in table 3. However, there also was a sizable number of claims addressed at the German government, but related to homeland issues (6% of all claims). For instance, Kurdish and Iranian groups accused the German government of supporting repression in their homelands by selling military equipment to the Turkish and Iranian regimes. In a way, such claims may be considered to be examples of transnational collective action, although of a rather different type than that which is assumed in the literature on postnational citizenship. While we found very few examples of migrants trying to mobilize supranational or homeland authorities to improve their position in the countries of immigration, we do find some support for transnational claims making in the other direction; that is, attempting to mobilize host society authorities or using the relative political freedom in the host society against the homeland regime. The point is that there is little new or postnational about such forms of transnational claims making by migrants, exiles, and political refugees. Just ask Karl Marx, Mikhail Bakunin, or the Ayatollah Khomeini.

Our main point regarding the greater importance of homeland-related organizations, addressees, and claims in the German context of an ethnic-exclusive model of migrant incorporation has now sufficiently been made. We will therefore now focus the remainder of our analysis exclusively on those claims related to immigration and ethnic relations in the host society. We first ask how important claims making by migrants and ethnic minori-

TABLE 5

Share of Ethnic Minorities in All Claims Making on
Immigration and Ethnic Relations in Germany and
Britain, 1990–95 (%)

|                                     | Germany | Britain |
|-------------------------------------|---------|---------|
| Migrants and ethnic minorities      | 6.8     | 20 7    |
| Other actors                        | 93.2    | 79.3    |
| Total                               | 100 0   | 100.0   |
| N                                   | 5,396   | 1,047   |

Note —$\chi^2 = 205 65$, $P = 0 0$, $df = 1$

ties is in this respect, compared to the claims making by actors based in
the majority population of the host society, such as state and party actors,
or labor unions, churches, and other civil society groups. Both the postna-
tional and the multicultural perspectives strongly emphasize the active
role of minority organizations in challenging immigration and minority
politics (hypothesis 4.2). Table 5 shows, however, that minorities have
been able to play such a role only in Britain. British ethnic minorities
were responsible for 21% of the total claims making on immigration and
ethnic relations, a higher percentage than was achieved, for instance, by
the national government (14%) or by members of the national parliament
(13%). With 7% of all claims, German minorities played a much less prom-
inent role in the public debate on immigration and ethnic relations, which
may have been about them, but was not nearly to the same extent as in
Britain influenced by them. At the risk of becoming repetitive, we again
have to conclude that the national citizenship model is best able to make
sense of these striking cross-national differences. In Britain, ethnic minori-
ties can challenge the official politics of ethnic relations because they com-
mand the political resources and are offered the opportunity to do so.
Britain's race relations and antidiscrimination legislation and its self-pro-
claimed status as a "multicultural society" constitute important discursive
resources for minorities, who can challenge the British state by referring
to its own aims and promises.

Finally, we focus our attention in closer detail on the contents of mi-
grants' claims on immigration and ethnic relations. Although our earlier
examinations of the model have not been very positive, we must still allow
for the possibility of more indirect forms of postnational claims making,
which do not directly involve transnational actors either as claims makers
or as addressees. To be fair, it has to be said that in their more prudent
formulations of the postnational thesis, authors such as Jacobson and Soy-

TABLE 6

TERRITORIAL FRAME OF REFERENCE OF MINORITY CLAIMS
MAKING ON IMMIGRATION AND ETHNIC RELATIONS IN
GERMANY AND BRITAIN, 1990–95 (%)

|                        | Germany | Britain |
|------------------------|---------|---------|
| Supra/transnational    | 1 0     | 1.4     |
| European               | 2       | · ·     |
| Bilateral              | .2      | 2.3     |
| National               | 98.6    | 94.4    |
| Unclassifiable         | ·       | 1 9     |
| Total                  | 100 0   | 100.0   |
| N                      | 421     | 213     |

NOTE —$\chi^2 = 15.77, P = 0\ 003, df = 4$

sal acknowledge that the nation-state in many cases remains the forum
through which postnational claims making has to pass. In this view, na-
tional organizations may make claims on national governments, but frame
their demands with reference to transnational or supranational actors,
rights, and conventions (hypothesis 4.3). This may include calls on na-
tional governments to take action on the supranational level (e.g., to strive
for a common European regulation of asylum procedures) or to respect
international conventions (i.e., the Geneva Convention on refugees). We
also include in this category claims that refer to the consequences that
national policies and developments may have on the transnational level.

In table 6 we show the territorial frame of reference of migrant claims
making. Even in this case, however, we find little support for the postna-
tional model, with 94% of claims made in Britain being firmly locked
within a national frame of reference, and as much as 99% of those in
Germany. Again Europe is surprisingly off the agenda in both countries.
In fact the only example we found was a statement by German Roma
and Sinti organizations calling on the German government to include the
Romani language in the European Charter on Regional and Minority
Languages. Other claims with a supra- or transnational dimension were
also rare, both in Britain (three claims by organizations invoking the su-
pranational authority of Islam) and in Germany (four claims by Jewish
groups backing their claims for combating xenophobia and anti-Semitism
by referring to the negative consequences that extreme right tendencies
in Germany might have for the country's image abroad and for foreign
investments in the German economy). In Britain, we find five claims
(2.3%) with a bilateral dimension, including for instance a black member

TABLE 7

CLAIMS BY MINORITY ACTORS FOR RIGHTS IN GERMANY AND BRITAIN, 1990–95 (%)

|  | Germany | Britain |
|---|---|---|
| Citizenship rights | 11.5 | .. |
| Other political and civil rights | 26.9 | 11.5 |
| Social and economic rights | 3.8 | 20.5 |
| Cultural rights | 46.2 | 50.0 |
| Antidiscrimination rights | . | 15 4 |
| Equal rights, unspecified | 11.5 | 2.6 |
| Total | 100 0 | 100.0 |
| N | 26 | 78 |
| Claims for rights as a percentage of all migrant claims making on ethnic relations (i.e., immigration and asylum issues excluded) | 8.9 | 42 4 |
| N | 291 | 184 |

NOTE.—$\chi^2 = 22\ 58$, $P = 0.0$, $df = 5$ for claims as listed in the top portion of the table $\chi^2 = 74.02$, $P = 0\ 0$, $df = 1$ for claims for rights as a percentage of all migrant claims making

of parliament demanding reparations payments to British blacks and their homelands as a compensation for British involvement in slavery.[31] All in all, however, the results of table 6 underline our previous findings that a few examples of claims making beyond national boundaries exist, but that they are the exception rather than the rule and cannot in any way be considered a representative characteristic of minority claims making. Given the fact that we have now given the postnational model three chances to prove its validity (with regard to actors, addressees, and the content and framing of claims) the clear conclusion to draw seems to be that from our data the postnational model of citizenship has little empirical underpinning, to the extent that it risks coming apart at the seams.

The last important feature of minority claims making that we wish to examine concerns the type of rights that are demanded. This is particularly relevant for the multicultural citizenship thesis, which sees demands for special group rights related to the recognition and protection of cultural differences—cultural rights—as central characteristics of minority claims making (hypothesis 4.4). In table 7 we present only those claims that are made in the field of ethnic relations (corresponding to the catego-

[31] In addition, as the table shows, there were four cases in Britain that were unclassifiable with regard to the territorial frame of reference of the claim. These were all made by Islamic groups whose claims consisted of asserting the importance of Islamic faith or the greatness of Allah, but without the contextual grounding that would enable their territorial categorization.

ries "minority integration politics" and "antiracism/xenophobia" in table 4) that make specific demands in terms of rights. The first point to make is that our findings give some credence to multicultural citizenship, as about half of the claims for rights in both countries are "cultural," with most of these being group specific.[32] In the German case, by far the most important group demanding cultural rights (five of 12 cases) were Roma and Sinti (gypsies) demanding recognition as a national minority, with the same rights as the Danish and Sorbian (a Slavic people in the Czech-Polish border region) minorities. Since Roma and Sinti have been traveling through Germany and the rest of Europe for centuries, they occupy a different position from that of recent immigrants such as the Turks and may therefore with some legitimacy claim the status of a national minority. In that sense, their important contribution to multicultural claims making in Germany confirms Kymlicka's (1995) point that national minorities have a stronger basis for claiming special cultural rights than ethnic groups of immigrant origin. Apart from this special case, more recent immigrants to Germany made very few claims to cultural rights, or, for that matter, to any rights at all. Claims for rights amounted to only 9% of all minority claims making on ethnic relations, the remaining 91% being mainly defensive claims against discrimination and xenophobic violence (see also table 4). The potential for societal disintegration as a result of culturally based demands seems thus very remote in the contemporary German context.

In contrast, we find more evidence to support the multicultural thesis in Britain. Here, the demands for specific group cultural rights were predominantly made by Muslims (23 of 37 cases), with others being under the self-definition of blacks, Jews, Rastafarians, African-Caribbeans, Indians, Pakistanis, and gypsies. These include examples of exactly the kind of cultural claims that are much discussed in the literature on multiculturalism. Thus the Muslim UK Action Front demands a cultural extension of British blasphemy law, in the words of its spokesperson: "The case will point out the common ground between Christianity, Judaism and Islam. In many past blasphemy cases . . . protection was afforded to the Old Testament, which was essentially Judaism."[33] We also have an example of a British headscarf case, where the activist states: "There was nothing

---

[32] Demands for special group rights (i.e., for a specific minority group) characterize most of the cases of cultural rights in both countries (10 of 12 in Germany; 37 of 39 in Britain). Most of the demands for rights in the other categories were claims for non-group-specific rights, such as equal citizenship and voting rights in Germany and equal rights in the labor market in Britain. In total 59% of the demands for rights in Britain were for specific group rights, compared to 38% of those in Germany.

[33] Note that our method allows analysis of the strategic action of claims in relation to the qualitative detail of the "speech" that is actually reported.

in the school rules about headscarves when we went there. The rules were amended last year. . . . We are not fanatics or fundamentalists. We just want the right to continue our education and practice our faith." What is notable from these examples, which are not atypical, is that the demands are made for an extension of the existing British multiculturalism. In the first case, the Muslims want to be added onto the list of acknowledged religions in the same way that they claim an earlier wave of migrants—the Jews—were incorporated. Likewise the headscarf claims are demands made in an inclusive rather than an incommensurable way. We have very few examples of more fundamentalist Islamic demands that challenge rather than address the authority of the nation-state. One made by a group in relation to Rushdie was, "Let us take him to Medina and let us stone him to death. Every Muslim should be prepared to cast the first stone."

It seems therefore that apart from the odd exception, the majority of cultural demands are better explained by the national context of a state-sponsored multiculturalism that gives opportunities for framing demands in this way. This argument is backed up by the comparative finding that there are many more demands for rights as a percentage of claims making on ethnic relations (42% vs. 9%), and claims making on ethnic relations is itself more important (table 4) in Britain than in Germany. That Muslim identity has difficulty fitting in the British context may be explained by the fact that British discrimination legislation does not acknowledge the right to religious equality (except in Northern Ireland) and so Muslims are in real terms offered only the rhetoric and not the substance of equality in British multiculturalism. Indeed of the special group rights claimed by blacks, more were framed in terms of civil or social rights than as cultural rights, one example being the right of blacks to be tried by juries with black representation. This returns us once more to the point we have made several times about the impact of the racial configuration of the British citizenship model on claims making.

CONCLUSIONS AND DISCUSSION

Our leading question in this article has been whether the claims making of migrants and ethnic minorities fundamentally challenge the liberal nation-state, as is supposed in theories of postnational and multicultural citizenship. By making an overview of the results of our comparative analysis of claims making by migrants and ethnic minorities in Britain and Germany in the 1990s, we find that the answer to this question must be negative. This is certainly the case with regard to the postnational model. Contrary to this perspective, which sees the nation-state as increasingly "insignificant" and "irrelevant" (Soysal 1998, pp. 208, 211) the nation-state continues to be by far the most important frame of reference for the identi-

ties, organizations, and claims of ethnic minorities, and national authorities remain the almost exclusive addressees of the demands of these minorities. To be sure, some of migrant claims making, especially in Germany, transcends national borders, but it does so in ways that are not postnational in any meaningful sense. These forms of claims making simply take another nation-state—not the host society, but the migrants' homeland—as their frame of reference. In some cases, such as that of the Kurds and the Kosovo Albanians in Germany, they even strive for the creation of new independent nation-states in their distant homeland. Such claims making by political exiles is not a new phenomenon at all but has accompanied the nation-state since its formation.

By contrast, we found precious little evidence for the types of claims making that are, according to the postnational model, typical for the modern migrant experience: transnational migrant organizations intervening in national politics, migrants addressing supranational institutions, minorities making demands on national governments in the name of international legal conventions and rights. None of these forms of postnational claims making accounted for more than at most a small percentage of the claims making of migrants and minorities. While this holds for postnational claims making generally, the almost complete absence of claims making related to the EU is a particularly surprising finding, which suggests that European integration has not progressed nearly as far as the rhetoric of both proponents and opponents of this project would have us believe. For the moment, Europe continues to be a coordinating committee for nation-states, laden with symbolism but with little autonomous power. Migrants in Europe seem to have understood this better than many social scientists, as have, by the way, European voters, the majority of whom do not even bother to vote for the Strasbourg puppet show called the European Parliament.

The last line of defense of postnationalists might be to argue that universal human rights declarations and conventions have been incorporated in the constitutions and legislation of liberal nation-states and that therefore even claims making that remains fully confined to the nation-state may have a postnational dimension. This watered-down version of postnationalism implies a diffusion of universal human rights principles from the supranational to the national level, a claim that is not supported by historical evidence. Human rights appeared as "self-evident truths" in national constitutions and bills of rights long before words such as "supranational" or "globalization" had been conceived. They are primary constituent principles of modern liberal nation-states, and are more of a basis for extending national citizenship than a pallbearer carrying it to an early grave.

Saskia Sassen (1998, p. 73) might have a point when she argues in de-

fense of the postnational position that it is often easier to point at continuities than to detect new, emergent developments. But even if we see postnationalism merely as an incipient trend and not as a state of affairs—which is certainly not the position of bold postnationalists such as Jacobson and Soysal—we think the empirical evidence should at least produce significant indications of the dawning of such a new era. Of course it is no problem at all to find examples that fit the postnational model; we have mentioned some in this article. However, they are rare exceptions rather than the rule of migrant claims making, and if they are indicative of a trend, it must be one that is at its earliest beginnings and remains for the moment more of a fiction than fact.

An additional weakness of the postnational model is that it stands empty-handed when it comes to explaining cross-national differences among liberal nation-states in reference to the claims making of migrants and ethnic minorities.[34] As our comparison of Germany and Britain has shown, cross-national differences in this respect are striking. While British immigrants mobilize on the basis of their racial and cultural difference from the majority society, the identities of German minorities refer to the national and ethnic categories of their homelands. Whereas homeland issues play an important role in Germany, they are marginal in the claims making of British minorities, who focus primarily on issues related to their rights and integration in British society. Insofar as German migrants focus on their position in Germany, their claims are primarily defensive and directed against overt xenophobia and racist violence. While such claims are not unimportant in Britain, too, we find in addition a sizable number of proactive claims for extensions of minority rights. Finally, we have shown that ethnic minorities in Britain play a much more important role in the overall public discourse on migration and ethnic relations than do their German counterparts, who as foreigners are not only symbolically but to an important extent also materially excluded from the political community.

This brings us to theories that stress the continuing relevance of national models of citizenship for the incorporation of migrants and ethnic minorities. Within the European context, Germany and Britain represent in many respects diametrically opposed incorporation regimes, which we

---

[34] To the defense of Yasemin Soysal it must be noted that in *Limits of Citizenship* (1994) she acknowledges and presents important evidence on cross-national differences in migrant and minority mobilization in different European countries. To explain these differences she points, much along the lines of the argument here, at the different traditions and institutional arrangements of migrant incorporation in these countries. In that sense Soysal's empirical analyses, which point at the importance of national contexts, are strangely at odds with her main theoretical conclusions, which emphasize the irrelevance of national compared to postnational forms of citizenship.

have labeled "ethnocultural exclusionist" and "multicultural pluralist," respectively. While the German model puts up important symbolic and material barriers to the acquisition of membership in the political community by foreign migrants, the British model provides for much easier access to citizenship, and in addition gives limited but still substantial recognition to their cultural difference. These liberal policies for resident migrants, it must be said, are coupled with immigration policies that make it in many respects more difficult for new migrants to enter the national territory than in the German case. Also in other respects, it is inappropriate to interpret German-British differences in a "good guys versus bad guys" frame. The British model of migrant incorporation is to an important degree a consequence of its imperial, colonial past, and by way of several restrictive changes in citizenship legislation subsequent British governments have done much to limit these consequences. By contrast, Germany has recently made an important step toward abandoning its ethnic conception of nationhood by way of a fundamental change in its citizenship legislation, which will come into force in 2000.

However, for the moment Germany and Britain continue to embody two markedly different models of nationhood and migrant incorporation. The hypotheses we have drawn from these two models fit our evidence on minority claims making remarkably well. Far from being insignificant, merely symbolic leftovers from a time when the nation-state still counted, national models of citizenship have important consequences for the identities, forms of organization, and types of claims of ethnic minorities. To an important extent, patterns of migrant claims making mirror the way in which the nation-state in the two countries defines the relation between ethnic minorities and the political community. Germany sees immigrants as foreigners, and that is exactly the way in which German minorities behave: they organize and identify themselves on the basis of their national origin and are still, in spite of residing in Germany several decades, preoccupied with the politics of their homelands. By contrast, the British multicultural state treats its immigrants as racial and cultural minorities within British society; again, that is pretty much how British migrants have come to see themselves: they mobilize as blacks or British Muslims and make claims on the British state for equal opportunity and multicultural rights.

In our theoretical discussion of national models of citizenship, we have identified a third ideal-typical model of citizenship, namely the assimilationist model, which in the European context is best represented by France. It will be an interesting question for further research to investigate whether a distinct pattern of migrant claims making can be identified in France, which matches its particular model of migrant incorporation. On the basis of the findings reported here, we may hypothesize that such a

French pattern of claims making will have certain similarities with our results for Britain. Because migrants have easy access to membership in the political community in France, too, we would expect them to have weak ties to their homelands' politics, as well. However, because of French pressures toward assimilation and the lack of recognition for ethnic and cultural difference, we would expect a different type of collective identity (more interethnic and based on the common status of immigrant) and different claims (equality and antiracism, but not a strong focus on cultural rights) than in Britain.

Extending the analysis to include other countries would also allow us to arrive at more conclusive answers about the challenge of multicultural citizenship. The evidence from Britain and Germany is mixed. In Germany, we found little evidence of strong pressures from ethnic and cultural minorities for the recognition of cultural identities and special group rights on the basis of cultural difference. In Britain, however, we found a substantial number of such claims, mostly, though not exclusively, originating in Britain's Muslim community. At the very least, the lack of a significant number of such demands in Germany, which also has a large Muslim population, shows that demands for multicultural rights are not a necessary consequence of the presence of a large population of non-European, culturally different migrants, but depend on the availability of political opportunities that legitimate and provide incentives for such claims making. It may nevertheless be that the British experience shows where Germany will be going once it has liberalized its citizenship legislation and German minorities focus on their rights and position in Germany instead of on the politics of their homelands. That is certainly what opponents of such a liberalization fear. The French experience might again provide an interesting alternative here. Since France invites migrants into the political community on the basis of equality, but to the exclusion of cultural difference, it could follow that migrant claims for multicultural rights are not nearly as important in this country as in Britain. If that is the case, the multicultural challenge would not be an assault on the unity of the nation-state imposed on it by the pressures of cultural difference inherent to immigration, but a response to an opportunity structure willingly created by some, but not all, nation-states. Far from being a mechanism to appease pressures arising inherently from immigration, state-sponsored models of multiculturalism may then turn out to stimulate such claims in the first place. And if indeed the multicultural appetite grows the more you eat, there may ultimately be something to the warnings against a self-sustaining fragmentation of the nation-state that opponents of multiculturalism fear.

However, even Britain does as yet not come close to such a scenario, and there is no indication that it will in the near future. Most demands

by minorities in Britain seem to be made within the national configuration of citizenship. Even the special group demands for cultural rights that are predominantly made by Muslims are perhaps best explained by the resistance of race relations policies to the recognition of religious equality. British multiculturalism offers Muslims a racial equality that contradicts their self-definition as a religious group. However, with the exception of the relatively few examples of incommensurable claims of fundamentalist Islamic activists, most Muslim claims demand an extension of the classical understanding of political and social equality in citizenship, to include the recognition of religious equality. Given the appropriate "cultural tools," British Muslims might be more likely to support than challenge the nation-state.

Although we are convinced that our analysis provides strong support for the importance of national models of citizenship and migrant incorporation, we do not want to conclude this article without having addressed a final caveat that may have been on some readers' lips for a while. There is an alternative explanation for the differences found between Britain and Germany that does not focus on national politics, but on the differences between individual migrant groups. As we have seen, Britain and Germany have drawn their immigrants from different countries and regions of the world. In Britain, the largest groups come from the Caribbean and the Indian subcontinent, in Germany by far the largest group of migrants come from Turkey. So perhaps the preoccupation of German claims making with homeland issues is just a peculiarity of Turkish and Kurdish migrants, and the focus on cultural issues a typical characteristic of Muslims from Pakistan and Bangladesh. This is a rather atheoretical and primordialist explanation, but still cannot be discarded out of hand. The question remains, in our view, why Turkish Muslims would be so different from Pakistani Muslims. Conversely, it cannot be maintained that the Indian subcontinent has no equivalent of the Turkish-Kurdish conflict that could fuel homeland-related claims making from this region in Britain. So the question is why Pakistani-Indian, Hindu-Sikh, and a myriad of other political conflicts on the Indian subcontinent do not to a significant degree translate into related claims making by British minorities from these regions. While we consider these to be important arguments against the alternative interpretation, the ultimate test would of course be to compare the claims making of migrants from the same region of origin in different countries of settlement. Because European immigration countries tend to have their own particular areas of migrant recruitment, this is easier said than done. Adding France to the comparison would not help much, for instance, because the most important French migrant groups come from the Maghreb countries and Portugal, which are marginal sources of immigration in both Britain and Germany. The

Netherlands, however, would be a helpful case in this regard, because it has sizable minorities from the Caribbean (allowing a comparison with Britain), from Turkey (which could be compared with German Turks), as well as from Morocco (comparable with France).

With our comparison of Britain and Germany, we have, to our knowledge, been the first to present a systematic empirical comparison of the claims making of migrants and ethnic minorities in two of the most important immigration countries of Western Europe. Our interpretation of the striking differences between these two countries in the framework of national models of citizenship and migrant incorporation has shown what merits such an approach may have. However, as our concluding remarks indicate, many important questions regarding the relationship between nation-states, immigration, and minority claims making remain. The challenge for further research therefore is to extend the kind of analysis we have presented here to include other countries and models of incorporation, both in Western Europe, and in the classical immigration countries of the New World.

REFERENCES

Bade, Klaus J. 1994. *Ausländer, Aussiedler, Asyl.* München: C. H. Beck.
Baldwin-Edwards, Martin, and Martin Schain, eds. 1994. *The Politics of Immigration in Western Europe.* London: Sage.
Bauböck, Rainer. 1994. *Transnational Citizenship: Membership and Rights in International Migration.* Aldershot: Edward Elgar.
Brubaker, Rogers. 1992. *Citizenship and Nationhood in France and Germany.* Cambridge, Mass.: Harvard University Press.
Castles, Stephen. 1995. "How Nation-States Respond to Immigration and Ethnic Diversity." *New Community* 21 (3): 293–308.
Castles, Stephen, and Mark Miller. 1993. *The Age of Migration: International Population Movements in the Modern World.* London: Macmillian.
Çinar, Dilek. 1994. *From Aliens to Citizens: A Comparative Analysis of Rules of Transition.* Wien: Institut für höhere Studien.
Cornelius, Wayne, Philip Martin, and James Hollifield, eds. 1994. *Controlling Immigration.* Stanford, Calif.: Stanford University Press.
Favell, Adrian. 1998. *Philosophies of Integration: Immigration and the Idea of Citizenship in France and Britain.* Houndmills, Basingstoke: Macmillan.
Forbes, Ian. 1995. "Institutionalising Anti-Discrimination in Europe." Pp. 192–208 in *Racism, Ethnicity and Politics in Contemporary Europe,* edited by Alec G. Hargreaves and Jeremy Leaman. Aldershot. Edward Elgar.
Freeman, Gary P. 1998. "The Decline of Sovereignty? Politics and Immigration Restriction in Liberal States." In *Challenge to the Nation-State: Immigration in Western Europe and the United States,* edited by Christian Joppke. Oxford: Oxford University Press.
Habermas, Jürgen. 1996. "The European Nation-State—Its Achievements and Its Limits: On the Past and Future of Sovereignty and Citizenship." Pp. 281–94 in *Mapping the Nation,* edited by Gopal Balakrishnan. London: Verso.
Hammar, Tomas. 1985. Introduction to *European Immigration Policy,* edited by Tomas Hammar. New York: Cambridge University Press.

Ireland, Patrick. 1994. *The Policy Challenge of Ethnic Diversity: Immigrant Politics in France and Switzerland*. Cambridge, Mass.: Harvard University Press.

Jacobson, David. 1996. *Rights across Borders: Immigration and the Decline of Citizenship*. Baltimore: Johns Hopkins University Press.

Joppke, Christian. 1996. "Multiculturalism and Immigration: A Comparison of the United States, Germany, and Great Britain." *Theory and Society* 25:449–500.

———. 1997. "Asylum and State Sovereignty. A Comparison of the United States, Germany, and Britain." *Comparative Political Studies* 30 (3): 259–98.

Kastoryano, Riva. 1996. *La France, l'Allemagne et leurs immigrés: Négocier l'identité*. Paris: Armand Colin.

Kleger, Heinz, and Gianni D'Amato. 1995 "Staatsbürgerschaft und Einbürgerung Oder: Wer ist ein Bürger? Ein Vergleich zwischen Deutschland, Frankreich und der Schweiz." *Journal für Sozialforschung* 35 (3/4): 259–99.

Koopmans, Ruud, and Hanspeter Kriesi 1997. "Citoyenneté, identité nationale et mobilisation de l'extrême droite. Une comparaison entre la France, l'Allemagne, les Pays-Bas et la Suisse." Pp. 273–94 in *Sociologie des nationalismes*, edited by Pierre Birnbaum. Paris: Presses Universitaires de France.

Koopmans, Ruud, and Paul Statham. 1999a. "Ethnic and Civic Conceptions of Nationhood and the Differential Success of the Extreme Right in Germany and Italy." Pp. 225–51 in *How Social Movements Matter*, edited by Marco G. Giugni, Doug McAdam, and Charles Tilly. Minneapolis: University of Minnesota Press.

———. 1999b. "Political Claims Analysis: Integrating Protest Event and Political Discourse Approaches." *Mobilization: The International Journal of Research and Theory about Social Movements, Protest and Collective Behavior* 4 (2): in press.

———. 1999c. "Political Claims-making against Racism and Discrimination in Britain and Germany." In *Comparative Perspectives on Racism*, edited by Jessika ter Wal and Maykel Verkuyten. Aldershot: Ashgate, in press.

———, eds. 2000. *Challenging the Politics of Ethnic Relations in Europe: Comparative and Transnational Perspectives*. Oxford Oxford University Press, in press.

Kriesi, Hanspeter, Ruud Koopmans, Jan Willem Duyvendak, and Marco G. Giugni. 1995. *New Social Movements in Western Europe: A Comparative Analysis*. Minneapolis: University of Minnesota Press.

Kymlicka, Will. 1995. *Multicultural Citizenship: A Liberal Theory of Minority Rights*. Oxford: Clarendon.

Layton-Henry, Zig 1994. "Britain: The Would Be Zero Immigration Country." In *Controlling Immigration: A Global Perspective*, edited by W. A. Cornelius, P. L. Martin, and J. F. Hollofield. Stanford, Calif.: Stanford University Press.

Lederer, Harald W 1997. *Migration und Integration in Zahlen*. Bamberg: Europäisches forum für Migrationsstudien.

Lie, John. 1995. "From International Migration to Transnational Diaspora." *Contemporary Sociology* 24:303–6.

Mason, David. 1995. *Race and Ethnicity in Modern Britain*. Oxford: Oxford University Press.

McAdam, Doug 1982. *Political Process and the Development of Black Insurgency, 1930–1970*. Chicago: University of Chicago Press.

McCarthy, John D., Clark McPhail, and Jackie Smith. 1996. "Images of Protest: Dimensions of Selection Bias in Media Coverage of Washington Demonstrations, 1982, 1991." *American Sociological Review* 61.478–99.

Modood, Tariq. 1988. " 'Black,' Racial Equality and Asian Identity." *New Community* 14 (3): 379–404.

Modood, Tariq, et al. 1997. *Ethnic Minorities in Britain: Diversity and Disadvantage*. London· Policy Studies Institute.

Münz Rainer, Seifert Wolfgang, and Ralf Ulrich. 1997. *Zuwanderung nach Deutschland: Strukturen, Wirkungen, Perspektiven*. Frankfurt: Campus.

Olzak, Susan. 1989. "Analysis of Events in Studies of Collective Action." *Annual Review of Sociology* 15:119–41.

Parekh, Bhikhu. 1996. "Minority Practices and Principles of Toleration." *International Migration Review* 30 (1): 251–84.

Rex, John. 1991. *Ethnic Identity and Ethnic Mobilisation in Britain.* Monographs in Ethnic Relations, no. 5. Warwick: Centre for Research in Ethnic Relations.

Rex, John, and Beatrice Drury, eds. 1994. *Ethnic Mobilisation in a Multi-Cultural Europe.* Aldershot: Avebury

Rucht, Dieter, Ruud Koopmans, and Friedhelm Neidhardt, eds. 1998. *Acts of Dissent: New Developments in the Study of Protest.* Berlin: Sigma.

Safran, William 1997. "Citizenship and Nationality in Democratic Systems: Approaches to Defining and Acquiring Membership in the Political Community." *International Political Science Review* 18 (3): 313–35.

Sassen, Saskia. 1998. "The De Facto Transnationalizing of Immigration Policy." In *Challenge to the Nation-State: Immigration in Western Europe and the United States,* edited by Christian Joppke Oxford: Oxford University Press.

Schlesinger, Arthur D. 1992. *The Disuniting of America: Reflections on a Multicultural Society.* New York: Norton.

Shain, Yossi, and Martin Sherman 1998. "Dynamics of Disintegration: Diaspora, Secession and the Paradox of Nation-States." *Nations and Nationalism* 4 (3): 321–46.

Shapiro, Gilbert, and John Markoff. 1998. *Revolutionary Demands: A Content Analysis of the Cahiers de Doléances of 1789.* Stanford, Calif.: Stanford University Press.

Smith, David M., and Maurice Blanc 1996· "Citizenship, Nationality and Ethnic Minorities in Three European Nations." *International Journal of Urban and Regional Research* 20 (1): 66–82

Solomos, John. 1993. *Race and Racism in Contemporary Britain,* 2d ed. London: Macmillan.

Solomos, John, and Les Back. 1995. *Race, Politics and Social Change.* London· Routledge.

SOPEMI (Continuous Reporting System on Migration of the Organisation for Economic Cooperation and Development). 1997. *Trends in International Migration: Continuous Reporting System on Migration.* Annual report 1996. Paris: Organisation for Economic Cooperation and Developement.

Soysal, Yasemin Nuhoglu. 1994. *Limits of Citizenship: Migrants and Postnational Membership in Europe.* Chicago: University of Chicago Press.

———. 1997. "Changing Parameters of Citizenship and Claims-Making· Organized Islam in European Public Spheres " *Theory and Society* 26:509–27.

———. 1998 "Toward a Postnational Model of Membership." Pp. 189–220 in *The Citizenship Debates: A Reader,* edited by Gershon Shafir. Minneapolis: University of Minnesota Press.

Tarrow, Sidney. 1989. *Democracy and Disorder: Protest and Politics in Italy, 1965–1975.* Oxford: Clarendon.

———. 1994. *Power in Movement: Social Movements, Collective Action and Politics.* Cambridge: Cambridge University Press.

Thränhardt, Dietrich. 1992. "Germany—an Undeclared Immigration Country." Pp. 167–94 in *Europe—a New Immigration Continent: Policies and Politics since 1945 in Comparative Perspective,* edited by Dietrich Thranhardt. Hamburg: Lit.

Young, Iris Marion. 1998. "Polity and Group Difference: A Critique of the Ideal of Universal Citizenship." Pp. 263–90 in *The Citizenship Debates: A Reader,* edited by Gershon Shafir. Minneapolis: University of Minnesota Press.

# Escape from the GDR, 1961–1989: Hybrid Exit Repertoires in a Disintegrating Leninist Regime[1]

Carol Mueller
*Arizona State University West*

Leading theorists contend that "exit" and its many manifestations do not fit the "contentious collective action" model. This article argues, however, that exit is frequently the only collective action available to average people in authoritarian regimes. In such regimes, collective exit may take a variety of forms, hybridized here from the elements of traditional and modern repertoires under varying conditions of opportunity and framing. Descriptions of hybrid exit episodes from the German Democratic Republic, 1961–89, are based on annual reports by Amnesty International, on secondary sources before 1989, and, primarily, on accounts from six international presses for 1989.

In a recent overview of an ambitious agenda for mapping "contentious collective action," McAdam, Tarrow, and Tilly (1996) acknowledge that their model may not apply to "shirking, flight, and what James Scott calls everyday forms of resistance" (p. 31). It is in fact the case that flight, escape, exile, exodus, and the search for asylum have rarely been treated as forms of contentious politics, collective action, or as variants of political process theory.[2] Why this is the case is less clear. What are the criteria by which some forms of collective claims making are excluded from the domain of contentious politics and others are included? What is it about collective flight against resistance that does not fit within the paradigm of contentious collective action?

[2] Work in the collective behavior paradigm has been much more likely to accommodate exit attempts (see Aguirre 1994; Snowden 1990; Miller 1985; Bach and Schram 1982).

Consider the most dramatic event in the demise of the German Democratic Republic (GDR), the "fall" of the Berlin Wall:

> *November 9, 1989. The Berlin Wall.* At the end of a long and rambling press conference, Gunter Schabowski, spokesman for the recently defunct Politburo of the East German Social Unity Party (SED), announced in an offhand manner that provisional travel regulations would be in effect until a new law was passed; namely, East Germans could now travel to the West without the usual restrictions on visas. Apparently, neither Shabowski nor the remaining Krenz government had intended to open the Berlin Wall, but East Germans who saw the press conference on television decided to see for themselves. Arriving at the checkpoints, crowds of East Berliners found that the exits were still barred and guarded as they has been for 28 years. Instead of going home to clarify the meaning of Schabowski's strange press conference the next day, they stood their ground shouting, "Open the gate! Open the gate!" to badly outnumbered armed guards. With television cameras feeding graphic images back to GDR audiences via West German stations, the standoff continued for three hours while the size of crowds continued to grow.
>  As taunts and shoving broke out at points of contact, the guards still had no instructions. Finally, at 10:30 P.M., the ranking East German border guards at Bornholmer Strasse and three other crossing points in the center of the city took matters into their own hands and opened the gates. Thirty minutes later, the Interior Minister ratified their decision with an official order By this time, one of the great celebrations of the century was underway at the Brandenburg Gate as tens of thousands poured through the checkpoints, and Berliners of East and West joined in toasting an historic moment. (This account is paraphrased from reports in the next-day issues of *Suddeutsche Zeitung, Frankfurter Allgemeiner Zeitung,* the *Times,* the *Guardian,* the *New York Times,* the *Washington Post;* see also Darnton 1991, pp. 11–12; Pond 1993, pp. 1–3.)

To be sure, there is much in the episode at the Berlin Wall that is atypical of the repertoires of collective action usually covered by theories of contentious politics or by political process theory. There is no organizational representation or formal presentation of claims; the crowd negotiates directly with representatives of the state through its numbers, body language, chants, and the unstructured conversations of those who happen to occupy positions of contact with border guards. The goal of exit into West Berlin is accomplished on the spot, not through protracted negotiations with multiple layers of state intermediaries. In most of its characteristics, the assemblies of crowds at checkpoints on the Wall resembled the premodern repertoire of collective action: *direct* in the selection of guards and gates as targets, *specific* in the claims made for immediate fulfillment, and *physical* in massing crowds to achieve the desired result (Tarrow 1994, pp. 39–45). This is not a demonstration or rally in which the media convey worthiness, unity, numbers, and commitment (Tilly

1995*b*, pp. 371–73) for the benefit of electoral blocs and intermediaries with the state. Yet, in their dependence on modern communication media to inform tens of thousands of East and West Berliners on where to assemble and the careful assurances to guards of nonviolent intentions, the crowds at the Wall shared characteristics of the modern repertoire. I argue that such a combination of modern and traditional repertoire elements is not unusual in collective actions constrained by the limits of authoritarian regimes. As described here, exit episodes created hybrid repertoires of collective action that drew on increasingly modern elements as the Leninist regime of the GDR slowly lost legitimacy and, finally, viability.

Increasing interest among sociologists in opposition, dissent, and resistance within authoritarian regimes (Pagnucco 1995, 1996; Aguirre 1995; Johnston 1991) suggests it is time for a reconsideration of the domain of theories that analyze collective action. The political process theory of collective action, with its roots in sociology, is the subject of interest here (but see Lichbach [1995] for a recent survey of alternative paradigms). A major tenet of contemporary political process theory has been that repertoires are created out of struggle in the context of specific political opportunities, forms of organization, and cultural frames (Tarrow 1998; McAdam, McCarthy, and Zald 1996; Tilly 1995*b*); thus application of the theory beyond democratic regimes is consistent with its basic assumptions.[3] In light of increasing application of political process theory to authoritarian regimes (Pagnucco 1995, 1996; Boudreau 1995; Brockett 1991, 1995), this discussion is also timely. The constraints of authoritarian regimes require the development of repertoires that exploit severely constricted political openings, resources, and opportunities for communication and assembly. The question addressed here is, How do these constraints influence the innovation of collective action repertoires in reference to a type of claim

---

[3] The present exploratory study of hybrid escape repertoires in the GDR over a 30-year period does not attempt to explain individual-level decision making encompassed by social psychological theories such as rational actor or rational choice. Repertoire theory is not inconsistent, however, with the type of rational actor theory used, for instance, by Karl-Dieter Opp and his colleagues (1995) in their study of individual decisions to participate in oppositional activity in Leipzig in 1989. Opp et al. work with a rational actor theory based on "soft incentives" and individual perceptions (pp. 31–32, 235) that incorporates macrolevel events, such as Hungary's decision to open its border and the mass exit of 1989, in terms of their effects on costs or benefits calculations of individual actors (pp 75–79). Rational action theory is interested in goals that can be defined as "public goods" (see Opp et al. 1995, p. 35; Lichbach 1995, p. 6) whereas repertoire theory applies to any collective good (or bad) whether it is to be shared with nonparticipants or not. This article traces the transformation of the exit goal over time to the point where it is redefined as a public good, the right to travel, in 1989.

and action (flight and escape) more characteristic of authoritarian than democratic regimes? In the historic circumstances of interest here, the constraints on collective action in the GDR disintegrated erratically over a period of 28 years after the construction of border fortifications in 1961 halted movement to the Federal Republic of Germany (FRG) and forced thousands to innovate new repertoires for the collective pursuit of exit.

I have selected the GDR as one example of the Leninist type of authoritarian regime that only a few years ago represented one-third of the world's population (White, Gardner, and Schopflin 1987, p. 3; Chirot 1986, p. 264) and now, as Jowitt points out, faces mass extinction (1992, pp. 249–283). The issue of present concern, exit as escape or flight, is one of the major types of collective action in such regimes. The GDR represents one among several Leninist regimes where exit was a particularly attractive alternative because of a shared border with a neighboring, more affluent state with a population of the same ethnic origin. Other examples of Leninist regimes that currently border states or population enclaves of the same ethnicity include North Korea, Cuba, and China, each with a somewhat unique configuration of factors influencing variable opportunities for escape and the appeal of the adjoining state.

The strategy of the analysis is to disaggregate traditional and modern repertoires of collective action into elements that can be recombined in an exploratory study of hybrid profiles. I use these elements to examine a series of exit episodes to determine what hybrid forms characterized the regime's move toward disintegration and allowed the expansion of opportunities for collective action. The analysis proceeds in three steps. The first section addresses three theoretical issues: "exit" is described as both a claim and a form of action with reference to Hirschman's work (1970, 1986) that increasingly has considered the constraints of authoritarian regimes; as a form of action, exit is then placed within the framework of Tilly and Tarrow's evolving theory of collective action repertoires; and, finally, the constraints on opportunities for action of Leninist regimes are described with particular reference to Eastern Europe and the GDR. In the second section, pressures for exit from the GDR, as well as the attraction of resettlement in the FRG, are indicated by a brief history of emigration from the close of World War II until 1989. In the final section, hybrid forms of collective action are described and illustrated with six exemplary exit episodes from 1961 through 1989. In conclusion, I characterize hybrid forms of action as one conceptual tool for expanding political process theory to encompass the constraints of authoritarian regimes. Alternative exit scenarios are suggested for several of the Leninist states mentioned above to indicate the highly adaptive nature of hybrid reper-

toires for realizing commonly held goals in conditions of limited opportunities through flight or escape.[4]

## EXIT AND REPERTOIRES IN LENINIST REGIMES
### Hirschman's Evolving Theory of Exit

Escape, flight, exodus, and the search for asylum are all types of behavior that demonstrate through the action itself varying levels of dissatisfaction with the conditions left behind and claims on some alternative life space. The most systematic comparison of this mode of expressing dissatisfaction with more familiar forms of protest is Albert Hirschman's (1970) classic statement, "exit, voice, and loyalty," where he argues that exit and voice represent alternatives available to customers, employees, and citizens who are displeased with the quality of goods and services they receive, the conditions of their employment, or the nature and effects of government policies. The two options represent the preferred mechanisms or "impersonations" (p. 15) of economics and political science, respectively, where exit represents the impersonal operations of the free market and voice, representing politics, is expressed through a wide range of actions from grumbling to revolution. In Hirschman's theory, exit is basically free, whereas voice requires creativity, effort, and organization. Thus, in his original formulation, exit would always be the preferred means of expressing dissatisfaction were it not for loyalty, which "holds exit at bay and activates voice" (p. 78).

In addition to their relative costs, Hirschman also tells us that exit and voice are inversely related; in particular, voice is the residual when exit is unavailable (pp. 33–36). Exit has the disadvantage, however, that once exercised, the opportunity for voice is lost, whereas the opposite is not true (p. 37). The remarkable juxtaposition of the two in the events of 1989 in the GDR led Hirschman (1993) to reconsider this "hydraulic effect" since it was the exodus of tens of thousands of East Germans that galvanized the mass demonstrations of fall. Together, the two mass movements brought down the regime.

The beauty of Hirschman's theory lies in its simplicity and its wide application (see, esp., Barry's [1974] review). Yet, as Hirschman recognized some years later (1986, p. 81), the theory has very little to say about the historically more frequent cases "where exit and voice are both in short

---

[4] The descriptions of exit episodes from the GDR are based on the annual reports of Amnesty International and secondary works for events before 1989; for the final crisis of 1989, I coded six international presses (Mueller 1997).

supply, in spite of many reasons for discontent and unhappiness" (see also Hirschman 1993; Laponce 1974).[5] In fact, Hirschman came to argue that both exit and voice depend on the expansion of democratic freedoms and have usually been restricted or employed together (1986, p. 79).[6]

To adapt Hirschman's formulation to authoritarian systems, exit will refer here to those collective actions designed to further personal and group interests by seeking either to leave a state against resistance or to secure the right to leave. Such actions encompass most of the forms of flight such as escape, search for asylum, and some forms of emigration where freedom of domicile and travel across geopolitical borders are legally prohibited and physically prevented. Because exit is defined here in terms of resistance, it can elicit riskier behavior in terms of personal costs than many forms of "voice." Exit actions, like other forms of contention, vary in terms of size and level of organization as well as those characteristics associated with traditional and modern repertoires of collective action (Tilly 1995a; Tarrow 1998). They may be aggregated (Marwell and Oliver 1984) to form more politically consequential "waves of emigration," "mass exodus," or "exit cycles" when multiple exit episodes coincide in time and space in a manner that can be meaningfully interpreted within a context of claims making.

As Hirschman (1986) recognized, all levels of exit action are disruptive and, therefore, threatening to an authoritarian regime, although the degree of threat varies according to most of the same factors that influence the threat capacity of "voice"—namely, the size, duration, and dramatic quality of actions, including violence. Symbolically, the flight of large numbers of people from a system, particularly against resistance entailing significant risk, discredits a regime and threatens its legitimacy as a functioning political system to a wider international audience and to those who stay. In addition, large-scale exit, like labor strikes in the contemporary repertoire of democratic states, deprives the regime of valuable human resources. It no longer serves as a "safety valve," but rather as a dangerous "loss of blood" (Hirschman 1986, p. 93).

Hirschman has made explicit the exit and voice options facing dissatisfied actors in democratic systems with open markets; and, as he pointed out later, in authoritarian systems where the costs of both options may be equally high. In such regimes, the question of interest, for present purposes, is not *whether* to exit, but *how*. These considerations encompass tactical and strategic choices: not only the specifics of how to get out at any given time, but how to adapt to authoritarian conditions the actions

[5] In reference to the Latin American experience, O'Donnell (1986) points out that both vertical and horizontal forms of voice are eliminated in authoritarian regimes.

[6] Others have contended that they may advance separately (Rokkan 1975; Finer 1974).

available in the repertoires of collective actions developed in the transition from traditional to modern times.

## Collective Action Repertoires

By the late 1990s, the tactics or activities for collective claims making are increasingly analyzed in terms of their place in a repertoire of collective action (Tilly 1978, 1979, 1984; Rucht 1990; Tarrow 1989, 1994, 1998; Traugott 1995; McAdam, McCarthy, and Zald 1996) or repertoire of contention (Tilly 1986, 1995a, 1995b; McAdam, Tarrow, and Tilly 1996). The concept of repertoires is regarded as one of the building blocks in a family of theories constructed over the last 25 years for explaining protest events, issue campaigns, social movements, and now rebellions and revolutions (McAdam, Tarrow, and Tilly 1996). The theories have been variously labeled as "resource mobilization" (McCarthy and Zald 1973, 1977; Zald and McCarthy 1987), "polity theory" (Tilly 1978), "political process theory" (McAdam 1982; McAdam, McCarthy, and Zald 1996), "collective action" (Tilly 1978, 1984; Tarrow 1989), and "contentious politics" (Tilly 1986, 1995b; McAdam, Tarrow, and Tilly 1996). Despite some differences, these theories share an interest in the explanation of claims making by subordinate or "relatively powerless" (Lipsky 1968) populations through nonelectoral collective action, although it is increasingly recognized that the two types of politics are not so different as was once thought (Meyer and Tarrow 1998). These theories also share the assumption that strategic decision making is one of the critical components in the social psychology of such joint actions (Klandermans 1984, 1997; Gamson 1992). Finally, it is assumed that these theories have universal import despite their origins in and primary application to Western democracies. For purposes of exposition, I will refer to the set of explanatory schemes as "political process theories" and to the generic activities encompassed as "collective claims making" or "collective actions" that assume variable forms as "collective action repertoires" or "repertoires of contentious politics."

There have been several recent attempts to expand the domain of political process theories to explain claims making in nondemocratic states (Pagnucco 1996, 1995; Boudreau 1996; Smith and Pagnucco 1992; Brockett 1991, 1995). Despite broad application to authoritarian regimes in Latin America, Eastern Europe, and Asia, these efforts have not included explicit analysis of the way in which claims-making activities are adapted or modified to meet the constraints on available opportunities and resources of authoritarian regimes. To further this effort, I adapt the concept of collective action repertoires developed by Charles Tilly and Sidney Tarrow for the analysis here. My examination will focus particular attention on the critical differences Tilly and Tarrow find between "traditional"

and "modern" forms of action. The hybrid escape forms discussed here incorporate features of both.

Originating in Tilly's (1978) model of mobilization, the basic idea of the repertoire is simply that the forms of contention available to a subordinate population for making collective claims are well defined and limited at any given time and place by that population's history of struggle, the available level of organization and resources, and the degree and type of political opportunity (Tilly 1979, p. 131). What is at stake in the theory of repertoires is the collective process of claims making or people acting together on the basis of shared interests in which "a limited set of routines are learned, shared, and acted out through a relatively deliberate process of choice" (Tilly 1995b, p. 41). They act on the basis of claims which, if realized, would affect the interests of other actors, leading to "contention" (p. 43).

The major empirical work in support of the theory has been the demonstration of changes in the repertoire in the United States, Britain, and France. Tilly's impressive studies of contention in early U.S. history (1979), in 400 hundred years in France (1986), and in almost a hundred years in Great Britain (1995b) have demonstrated that repertoires in these countries changed slowly over centuries. As states and markets expanded, politics became national in scope, and repertoires of contention changed to meet the new conditions.

The primary conceptual device for characterizing this change has been a typology, defined variously, of "traditional" and "modern" repertoires.[7] The traditional repertoire, characterizing collective actions for Britain, France, and the United States from approximately 1650 until 1850 (Tilly 1983), was based on the interactions of corporate and community groups contesting their interests in relatively isolated, local communities. Pulling down the houses of tax collectors and officials who had violated local norms of justice, confiscating grain and setting prices consistent with the local moral economy in times of scarcity, and breaking down enclosures of fields and forests traditionally allotted for community grazing and wood gathering were all directed at specific individuals, or their property, with whom two or more individuals had a grievance. As Tarrow points out in quoting the French historian, Marc Bloch, there was a strong connection between the grievance and the collective action employed to make a claim

---

[7] In a review of the corpus of Tilly's work, Tarrow has said, "The concept of the repertoire is the most ambitious attempt we have to date to trace historically the changes in the character of collective action in connection with changes in society and politics" (1996, p. 592). Tilly credits the origins of the distinction to his dissatisfaction with the typology of claims from his work in the 1970s that was mistakenly associated with theories of modernization (see Tilly 1995a).

(1993, pp. 75–76). Tarrow characterizes the traditional repertoire as parochial, direct, and particular (1998, p. 33), ideally suited to the informal modes of communication and association among known neighbors, kin, and coworkers of premodern societies (see also Tilly 1995a). Because of its parochial and specific focus, the traditional repertoire could not be transported beyond the local community or joined with collective actions elsewhere to create a national-level campaign or social movement (Tarrow 1993, p. 74).

The defining characteristics of the modern repertoire—cosmopolitan, autonomous, and modular (Tilly 1995b, p. 46)—require means of communicating beyond the local community and modes of association and coordination of action with individuals who are not necessarily known to each other (Tarrow 1998, pp. 42–53). Learned forms of acting collectively can be used in a variety of settings for a variety of claims, which gives them a modular or flexible quality uncharacteristic of traditional claims making (Tarrow 1993, 1998). The petition, demonstration, public meeting, and strike as well as the barricade (Traugott 1995) and the popular or urban insurrection (Sewell 1990; Piven and Cloward 1992; Tarrow 1998, pp. 39–40) can be adapted to myriad purposes.[8] Unlike the close connection between grievance and form of action characteristic of the traditional repertoire, modern forms of collective action are tied neither to specific claims nor to claimants.

And, how do such monumental shifts occur in the means of making collective claims? Both long-term and short-term answers have been proposed. In the long term, Tilly argues that it is an indirect result of macrolevel changes that affect political opportunities, forms of organization, and collective identities. In the cases he has studied in Western Europe, it was the expansion and encroachment on people's everyday lives of capitalist markets and the war-making state that led to the particular innovations that now characterize the modular modern repertoire. In different historic contexts, other macrolevel changes, such as "population movements, the disintegration of empires, religious revival, or environmental degradation" (1995b, p. 383) might have very different consequences for political opportunities, organization, identities, and, thus, the components of the repertoire.

In the short term, the process of innovation by which repertoires change occurs "chiefly at the edges of existing performances, as someone stretches

---

[8] To this list, Rucht (1990) would add a host of strategies and tactics from new social movements, and, finally, in a rare moment, Tilly (1983) would include sit-ins, civil disobedience, hijackings, and terrorist bombings. Steinberg (1995) even suggests that the discourse of mobilization be added to our understanding of the collective action repertoire.

known routines past familiar limits" (Tilly 1995*b*, p. 381). Such innovations, in Tilly's view, are designed not so much to catch authorities off guard as to mobilize more supporters to their cause. Similarly, Tarrow associates paradigmatic shifts in repertoires with peak periods of mobilization in protest cycles when heightened competition places emphasis on devising new tactics for attracting resources (1989, pp. 78–80; 1998, pp. 102–3). McAdam (1982) places more emphasis on the strategic importance of tactical innovations as part of an ongoing struggle with the state, or, in the case of the Civil Rights movement, with its countermovement. Those innovations that are adapted widely and become part of the modular repertoire are associated with the conspicuous success of a claimsmaker (Tarrow 1993).

Based on the features of developing, capitalist societies in North America and Western Europe, the modular repertoire described by Tilly and Tarrow has evolved with democratic polities based in the freedom and means of association and mass communication. It has gained its political influence by providing opportunities to demonstrate collectively the *worthiness, unity, numbers, and commitment* of people with grievances that have implications for electoral outcomes (Tilly 1995*b*, pp. 371–73). In concert with the spread of suffrage, the development of the modern repertoire over the last 200 years has been a major factor in establishing the culture and structure of democratic participation.[9]

For purposes of further comparison, it is useful to identify the elements or dimensions of analysis in this critical transition. Any specific form of action can be characterized in terms of its specific profile of elements. I have schematized the differences between modern and traditional repertoires in terms of their key elements for purposes of profiling several dis-

---

[9] Although a brief summary can hardly do justice to a theory elaborated over a period of 20 years by several of our most prolific scholars of collective action, it must be admitted that the differences between the "traditional" and "modern" repertoires are ever-evolving from the dialogue between Tilly, Tarrow, and their critics. This is particularly true for issues related to the associational base of the modern repertoire. Sewell (1990) has argued that Tilly's *Contentious French* (1986) fails to appreciate the decisive importance of the French Revolution for redefining the legal status and, therefore, privileges, of corporate actors who had previously made claims and become mobilized on the basis of their corporate identities. According to Sewell, the special purpose association did not become the predominant organizational actor in French collective action until the 1860s after a contentious century in which it was in or out of favor (legal or illegal) depending on the regime in power. In response, Tarrow (1993) points out that special interest associations appeared in both the English colonies and Britain without benefit of the French Revolution and that this type of association is not a prerequisite for those modern collective actions which gave rise to the social movement. Tilly's book on the contentious British (1995*b*) places the ascendance of the special interest association over communal and corporate claims makers by the 1830s.

tinctive hybrid combinations of exit actions in the GDR from 1961 through 1989. The defining elements of the two repertoire types are indicated in columns 1 and 3 of table 1. The key elements for describing the differences are these: (1) the composition of collective actors, (2) the criteria for the division of labor, (3) sites of activity, (4) the predominance of public activities in both types of repertoires, although some clandestine activity exists in the traditional repertoire, (5) local and primary modes of communicating versus modern reliance on print and electronic media, (6) goals of action that are concrete and specific to the immediate actors and their grievance versus more universal claims in the name of general classes of actors, (7) actions that are direct, personal, and often violent in contrast to the symbolic demonstration of "worthiness, unity, numbers and commitment," and, finally, (8) targets of action that are considered the direct source of the grievance versus actions directed to the media as a way of reaching third parties, including the electorate and state representatives.

Like Hirschman's initial work on exit and voice, the characteristics of the modern repertoire, as well as the conditions that facilitate its use, are uniquely suited to relatively open societies in which the costs to the individual of acting with others on a collectively shared grievance is largely a question of time and inconvenience (Oliver and Marwell 1992). The *public* performances of claimsmakers (Tilly 1995*b*, p. 43) presumes that the collective display of grievances is at least tolerated by the regime. Although there were state restrictions on both print and association well into the 19th century, particularly in France (Sewell 1990), the social foundations of the modern repertoire that depend on open communications and association also require rudimentary democratic conditions. Yet, it is well known that neither claims making, access to the media, nor free association can be taken for granted in any type of authoritarian regime, particularly those regimes variously labeled "totalitarian" (Arendt 1951) or the "Leninist" variant considered here (Jowitt 1978, 1992; Chirot 1991, 1986). To expand the domain of the theory to such regimes requires a consideration of the particular constraints they impose on collective action.

## The GDR as a Leninist Regime

The ideal-typical Leninist regime includes a combat-oriented communist party that claims absolute power to achieve an advanced, egalitarian utopia based on a charismatic authority justified by the scientific laws of Marxist historical development (Jowitt 1978; Chirot 1986). Although a diversity of conflicting interests exist in such regimes, their most important features for present purposes are the near-monopoly of power and authority by a vanguard party, the absence of a public arena in which individual

TABLE 1

TYPICAL PROFILES OF COLLECTIVE ACTION BY REGIME TYPE

| ELEMENTS OF ANALYSIS | REGIME TYPE | | | |
|---|---|---|---|---|
| | Traditional Authoritarian* | Modern Leninist Authoritarian† | Modern Democratic‡ |
| 1. Composition of collective actors | Corporate groups based on kinship and locality | Corporate groups based on kinship and locality in East, Semi-professional "helpers" assisted by volunteers in West; Criminal syndicates | Social movement organizations led by professional organizers; Paper memberships or local grassroots networks |
| 2. Division of labor/level of discipline | Simple, based on local social relationships | Complex, detailed, highly disciplined | Complex, disciplined in national offices, More informal at local levels |
| 3. Sites of activity | Local to residence | Local escape routes | Local, state, national, and global |
| 4. Public-private realm | Primarily public, Numerous exceptions such as poaching | Clandestine | Public |
| 5. Medium of communication | Primarily local, informal networks | Carefully monitored for secrecy except for fund-raising efforts in Western media | Highly dependent on mass media for framing appeals to third parties and potential participants |
| 6. Goals of action | Immediate, concrete, and specific to local situation and corporate actors | Immediate, concrete, and specific to corporate actors | General, universal claims on the state for public goods coupled with personal grievances |
| 7. Type of action | Immediate, concrete, and specific; Physical, often violent | Immediate, concrete, and specific to escape routes; Physical; Rare use of violence | Symbolic action messages framed in terms of electoral alignments |
| 8. Target of action | Local elites or representatives of growing states and markets | Barriers to escape | Representatives of the state at local, state, national, and global levels |

* Based on the work of Tilly (1978, 1986, 1995b) and Tarrow (1989, 1994)
† Based on GDR, 1960s early escape hybrid.
‡ Based also on the work of McCarthy and Zald (1973, 1977) and Morris (1984)

(as opposed to collective) citizenship rights obtain, and the illegitimacy of opposition to the "leading role" of the party. Given these constraints, there are limited opportunities for collective action because resources for mobilizing supporters and third parties are monopolized by the party. Open communication media, organizations, public gatherings, and collective identities that might compete with the party's leading role are prohibited. Universal human rights claims against the state are preempted by rights proclaimed and enforced by the revolutionary party in the name of the working class (Schopflin 1993, pp. 75–103; White et al. 1987, pp. 223–52; Lane 1984).

Despite their common features, Leninist regimes differ from each other and over time in terms of the unique historical contexts in which they arose (indigenous revolution or foreign occupation), their client or autonomous status, and the priorities assigned by the party to tasks associated with the mandates of Marxism-Leninism. Whether Leninist regimes are typologized in terms of developmental stages or combinations of regime priorities, the level of repression is a defining characteristic.[10]

The two schemes for differentiating Leninist regimes arrive at similar conclusions for Eastern Europe regarding the relaxation of repression and moderate openness to the outside world following Stalin's death in 1953. After the post–World War II takeover period, elite focus in Eastern Europe was on atomizing all sources of resistance in civil society (Schopflin 1993, pp. 89–93; White et al. 1987, pp. 237–38).[11] As indigenous opportuni-

---

[10] Scholars of Leninism who have distinguished regimes in terms of their developmental stages base these differences on a sequence of tasks endorsed by the party elite. Dallin and Breslauer (1970) identify three stages: takeover, mobilization, and postmobilization. Similarly, Jowitt (1978) argues for three stages: transformation of the old society, consolidation of the new regime, and, finally, inclusion of the nonapparatchiki sectors of society, e.g., Khruschev's "party of the whole people." It is only in the last stage that the opportunities for collective action increase substantially as the security apparatus is brought under control by the party elite. For the Soviet Union and Eastern Europe, this was the period following Stalin's death in 1953 This is the stage in which the modern repertoire was tested repeatedly through high-risk trial and error as elites tried to substitute manipulation and persuasion for more extreme forms of repression (Dallin and Breslauer 1970, p. 96). Chirot (1986, pp. 266–70) differentiates regimes in terms of five dimensions of elite priorities: equality vs. efficiency, development of heavy industry vs. light consumer goods and services, centralization or decentralization of economic control, the extent of openness to the outside world through control of the media and the border, and, finally, level of repression. After the mid-1950s he characterizes most of Eastern Europe, excluding Albania and Romania, in terms of a "liberalized Stalinist" model with "medium" levels of commitment on each of the five priorities. Wider coverage of theoretical models analyzing communist states can be found in Thomas (1997) and White et al. (1987, pp. 14–30).

[11] For instance, the number of Hungarian interest-group clubs not connected to sports declined from 14,000 in the early 1940s to fewer than 1,000 after the takeover (Hankiss 1990, p. 20). Opportunities for pubic association outside the party were stripped away.

ties for association were eliminated, so also were the opportunities for developing the overlapping networks that are critical for mass mobilizations (Oberschall 1973) and for creating the "constituted groups" that develop and maintain repertoires of collective action (Tilly 1979, p. 132). During this period the secret police in Eastern Europe, like their predecessors in the Soviet Union, enjoyed their greatest autonomy. There were few visible critics, and these were dealt with summarily along with many who had voiced no criticism at all (Holmes 1986, p. 275). Their power was dramatically demonstrated in the show trials of 1949–54 directed from Moscow and imitative of the Soviet purges of the 1930s (Schopflin 1993, pp. 96–97; Holmes 1986, p. 66). Although capital punishment was largely avoided in the GDR, one of every 200 adults served time in prison or internment camps (Maier 1997, p. 20).

With the takeover by the SED in 1948–49, East Germany, even more than other East European countries, became what Gunter Gaus termed a "niche" society in which each individual or family tried to withdraw into its private life (Maier 1997, pp. 28–30; Ash 1994, p. 194).[12] Public life required a higher level of outward conformity than even neighboring Czechoslovakia. This atomization of life was supported by an impressive State Security Service (*Staatssicherheitsdienst, SSD,* or *Stasi*) sometimes referred to by the opposition as "the firm" and by the West German press as "the octopus."[13] Surveillance extended to all forms of communications. Post-1989 examination of the Stasi files has revealed that no matter was too small for their attention–from domestic arguments to hairstyles and television programs. This remarkable penetration of civil society was accomplished with a staff of over 100,000 full-time employees and at least 300,000 informers (Jarausch 1994, p. 35). To live any kind of public life, it was difficult to avoid explicit or tacit collaboration (Maier 1997, pp. 149–53).

Opportunities for free association and open communication are always problematic in Leninist regimes, but the Soviet leadership struggle, re-

"All legally recognized organizations, regardless of their aim must submit to party control" (White et al. 1987, p. 247)

[12] Gaus served as the first permanent representative of West Germany in the GDR from 1974 until 1981. He is quoted by Maier (1997) as saying, "What is a niche in the society of the GDR? It is the preferred place for people over there, the place in which the politicians, planners, propagandists, the collective, the great goal, the cultural legacy—in which all of these depart so that a good man, with his family, among friends, can water his potted flowers, wash his car, play Skat, have conversations, celebrate holidays" (p. 29)

[13] Even as late as 1989, the Stasi had a communications center in Leipzig from which they could listen to every telephone conversation in the city, 2,000 simultaneously (Darnton 1991, p. 130).

sulting in Krushchev's policy of destalinization, marked the beginning of a period of experimentation such that opposition forces in each country tested the limits of the possible in a "dialectic of dissent and repression" (Sharlet 1978).[14] With the breakdown in elite consensus, regime opponents could try to recreate the infrastructures of association and communication that had been eviscerated in the takeover following World War II. Dramatic mass mobilizations were anticipated by the advent of new discussion groups.[15] As a prelude to broader participation, independent means of communication were created with discussion groups and workers councils publishing critical journals and newspapers that documented regime failures, articulated and legitimated collective grievances, and publicized manifestoes and lists of demands. Such sporadic and openly critical media were suppressed and then replaced, in the wake of Prague Spring, by a more institutionalized form of clandestine communication, *samizdat*, or uncensored and, therefore, unofficial publication (Schopflin 1993, p. 167; Skilling 1983; White et al. 1987, pp. 244–53). In Poland, where the dialectic of mass action and reform was most successful, the party's monopoly of the mass media had essentially disappeared by 1980 as Solidarity and the local unions established their own autonomous media (Ash 1990, pp. 59–60; Hirszowicz 1986, p. 178). After 1970, the Lutheran church in the GDR, like the Catholic church in Poland, was permitted to act as a semi-autonomous space in which opposition could be guardedly expressed and discussed.

With the partial thaw of détente in the 1970s, grassroots mobilization

[14] The "dialectic" proceeded erratically after Stalin's death as party leaderships in each of the East European regimes responded to local conditions in interpreting the Kremlin's demand for a "New Course" (Sharlet 1978). Throughout the region the post-Stalinist period was marked by a series of crises in which political opportunities were generated by competition and conflict among party elites, notably between hardliners and reformers Occasional incidents such as an author banned from the Writer's Union or sent into exile gave public indications of undercurrents of dissent among intellectuals. The most dramatic form of opposition until 1989 was the rare eruption of mass "unrest" in the form of urban insurrections or strikes (Lane 1984; Childs 1988, pp. 31–33; Holmes 1986). The major uprisings in Eastern Europe are well known: the 1953 workers' insurrection in the GDR, the Hungarian Revolution of 1956, the 1968 Prague Spring, and the series of Polish strikes in 1970–71, 1976, and 1980–81. Most of the popular efforts in these events reflected modifications in the modern repertoire—strikes, demonstrations, and urban insurrections—despite the absence of a meaningful electoral system in the countries where they occurred.

[15] In 1956, it was the Petofi Circle in Hungary, later suppressed; in 1968 Czechoslovakia, KAN (Club of the Non-Party Committed) and K231 (club of former Stalinist prisoners, later suppressed and then reinvented as Charter 77 in 1977); and, in Poland, KOR (Committee for the Defense of Workers), which reflected a new merger of intellectuals and workers in 1976, but was superseded by Solidarity in 1980 (Schopflin 1993, pp. 182–85; White et al 1987, pp 247–51).

began in the GDR as tiny groups within the population began to mobilize under the protection of the Lutheran church on issues of antimilitarism, peace, the environment, and, finally, gender (Allen 1991; Di Caprio 1990; Pierard 1990; Lemke 1986; Mushaben 1984). These efforts were devoted to the reform of socialism, not its overthrow. Though stunted by the constraints on civil society, the groups began to develop the first prerequisites of a modern collective action repertoire. They had open, though unpublicized, meetings at churches with sympathetic pastors where like-minded individuals could meet, despite the ever-present Stasi informers. When reform interests, like disarmament, coincided with goals of the state, such as the pursuit of superpower disarmament during the early 1980s, some church-related groups were permitted to form links with peace groups in the West (Ramet 1984; Woods 1986).

Outside the protection of the church, local groups had little knowledge of or communication with each other until the end of the 1980s (Philipsen 1993). Most public appeals that could have put them in touch with each other or a broader constituency would have led to repression. To overcome this isolation and to reach a larger audience, one of the most daring strategies for dissidents was to publish political statements in West Germany that would then be seen on television in most of East Germany or to make a political point at state occasions attended by the Western press (White et al. 1987, p. 245). Such actions almost inevitably led to arrest and expulsion. These sanctions were widely understood given the serious consequences that befell even famed writers and scholars, such as those who signed the petition for exiled folksinger, Wolf Biermann, in the late 1970s.

These, however, were rare forms of publicly expressed grievances. While undoubtedly the most common form of opposition was the retreat of the individual or family into private life, many expressed a more comprehensive level of dissatisfaction with the regime and their lack of hope for the future by leaving. For those whose dissatisfaction led them to reject the socialist experiment entirely and to disavow "loyalty" to friends, neighbors, and kin, there was a relatively straightforward solution until 1961— emigration to the FRG, where they would receive automatic citizenship, welfare benefits, job referrals, and housing preferences in an increasingly affluent country of shared language, culture, and history. But, as the border was gradually closed off, exit became the most costly option for opponents of the regime. After 1961, those who applied for emigration to the FRG faced the loss of their jobs, their apartments, their possessions, and, quite possibly, their friends and associates. By the late 1980s, applying for a permit too frequently or escaping were the major "crimes" of those serving time in prison for political reasons (Amnesty International 1989). It is this history that demonstrates the grassroots form of mass dissatisfac-

tion with life in the GDR. The story of exit is summarized here to provide the general outline of emigration before turning to the individual episodes that reveal the patterning of hybrid repertoires.

## EXIT AS DISSATISFACTION WITH THE GDR REGIME

In the years immediately following World War II, and before the western border of the GDR was closed by the Berlin Wall in 1961, political consolidation by a newly triumphant SED led to large-scale departures by their erstwhile allies in the Socialist Party (the SPD), followed shortly by supporters of the Christian Democratic Union (the CDU) and the Liberal Democrats (Naimark 1992, pp. 76–77), who fled after hundreds of their leaders were arrested, imprisoned, and, in some cases, "disappeared" to the Soviet Union (Childs 1988, p. 22). These overt attempts to find asylum in the other Germany were followed by further political pressures on the labor market of the GDR as one sector of the economy after another was brought under centralized control. In 1948, the textile industry and thousands of small factory owners and employees left Thuringia and Saxony; in 1949, it was commerce and trade; in 1950–51, nationalization of the health care system led doctors, dentists, and druggists to leave; and, in 1952, 20,000 farmers fled as the collectivization of agriculture got underway (see fig. 1).

Most departures combined political with economic interests. When the SED leadership announced at the second party conference of June 1952 that the second five-year plan would follow the Stalinist program of "building socialism" through the large-scale development of heavy industry financed by higher levels of austerity, every sector of society was burdened with higher taxes and insurance premiums. Preferences and discounts on consumer goods and ration cards for food were to be withdrawn from the middle class, while higher production quotas were established for workers (Childs 1988, pp. 31–33). These draconian measures not only precipitated the famous Workers Rebellion of June 1953 but also led to over 300,000 departures for West Germany, an exodus unequaled until 1989 (Torpey 1995, p. 23). Thus, "labor market pressures" resulted from the steady incorporation of one sector of society after another into a centralized, planned economy, not the opportunistic individualism that economists associate with free labor markets.

The installation of heavily armed fortifications on the Western boundaries of the GDR, culminating in construction of the Berlin Wall in August 1961, was meant to stop the mass exodus. The fortifications were highly successful. The annual average of out-migrants dropped from 210,000 before the wall was built to 21,000 afterward, a tenfold decline. A fortified border could now be used by the GDR as a threat both to those who
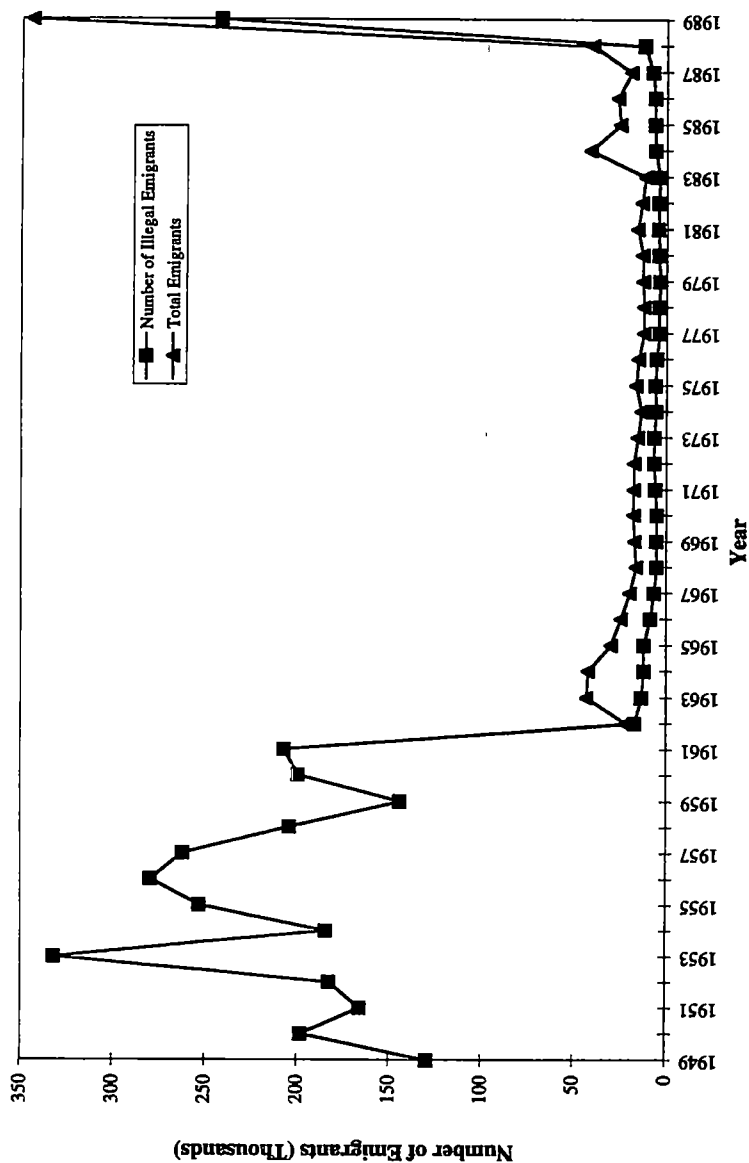
FIG. 1.—Emigration from the GDR to the FRG, 1949–89 (data from Wendt [1991, p. 390])

sought to leave for West Germany and to those who wanted to stay and seek reform. Prospective emigrants could be threatened with sanctions ranging from loss of jobs and friends to imprisonment and death, while reformers could be intimidated or silenced by the prospect of expulsion and exile (Ash 1993; Hirschman 1993).

In the period 1961–89, both tactics were perfected and others were invented by the regime in its attempts to control both expressions of dissatisfaction. The opportunity to leave the GDR legally was guaranteed by the 1975 Final Act of the Helsinki Conference on Security and Cooperation in Europe, as well as the UN Declaration of Human Rights and the International Covenant on Civil and Political Rights, all of which had been signed by official representatives of the GDR (Thomas 1997; Woods 1986, p. 32). Despite its status as a signatory of these international human rights documents, the GDR constitution did not guarantee its citizens the right to leave. Instead, to leave without permission was a violation of Article 213 of the Penal Code, as was solicitation of information or help from international organizations (Amnesty International 1989, p. 8). Penalties were more severe if individuals planned or acted together.

Unless the applicant was a pensioner or sought to deal with "urgent family matters" (birth, marriage, anniversary, or death), the process of application was designed to be demeaning, frustrating, and, usually, unsuccessful (Jarausch 1994, p. 18). Regulations were difficult to follow and changed frequently; applicants were forced to give up their passport, their real estate and other private property, as well as their prospects for meaningful employment. As they waited for permission, they were likely to be assigned onerous employment, such as cleaning toilets, and to live on the fringes of society, abandoned by friends and associates who feared contamination by association with subversives. Those who applied repeatedly for permission to leave were regarded as negatively as were those who sought to cross the border illegally.

There were only three ways to leave aside from urgent family matters: one could go legally by applying for and receiving a permit; one could leave illegally by finding a way to cross the border without detection; and one could leave in exchange for hard currency or valued goods from the West German government. Despite the obstacles, many did file applications to leave and sometimes found their chances influenced unpredictably by shifts in the political climate. When the Helsinki Accords—which guaranteed basic human rights such as freedom of movement—were signed by the GDR in 1975, hundreds of thousands filed applications to leave. Instead of receiving permission, they were treated as enemies of the state; no increase in the number of permits was granted (Thomas 1997; Ash 1993, p. 195). When the international climate changed again, appli-

cants might suddenly find they were successful. In 1984, for instance, Honecker was finally successful in visiting West Germany, thus achieving long-sought recognition for himself and legitimacy for the GDR. With tensions thus relaxed, five times as many applications were approved in 1984 as was the case in 1983 (Naimark 1992, p. 79).

Because the application process was arduous and uncertain, many tried to leave illegally. Although there is little documentation of their experiences, by the end of 1988 over 173,000 had reached the FRG without permission, their numbers increasing dramatically at the end of the 1980s (see fig. 1). Fewer than 200 were killed close enough to the border for Western sources to document their escape attempts, but capture guaranteed a prison term and other hardships (Amnesty International 1989, p. 8). Of the estimated 17,000 East Germans sentenced to prison terms between 1961 and 1984 for political reasons, the majority were charged with "fleeing" the republic (Naimark 1992, pp. 77–79). In Amnesty International's *Annual Reports* from 1974 through 1989, the vast majority of known political prisoners in the GDR were those who had sought to escape.

The fortified border served to enforce the threat of exile for those who sought to stay and reform what dissident Rudolf Bahro had satirized in 1977 as "really existing socialism." Although Bahro apparently left willingly after 10 months in prison for publication of his critique of socialism, other dissidents were not given this choice. Wolf Biermann, the balladeer whose biting lyrics had led the regime to prohibit his public performances, was deprived of his citizenship in 1976 while he was on a tour in West Germany. An unprecedented public petition for reconsideration signed by leading intellectuals was unsuccessful, and he remained in exile (Joppke 1995, pp. 66–69). Ten years later, the leaders of the Initiative for Peace and Human Rights (IFM) were given the choice of long prison terms or exile after the Stasi attacked and badly damaged their presses and confiscated copies of an underground newsletter. Such forced exile was one of the major forms of repression in Eastern Europe and the Soviet Union during the post-Stalinist period (Sharlet 1978). It was reserved for dissidents whose prominence made more subtle forms of administrative harassment and hooliganism ineffective—and for persistent emigration applicants (Sharlet 1978, pp. 766–67).

The final means of leaving was the exchange of GDR prisoners for hard currency (deutsche marks). Beginning in 1963, the FRG began offering deutsche marks or needed supplies like oil, copper, silver, or industrial diamonds to the GDR in exchange for prisoners (Ash 1993, p. 143). Most of those who were "ransomed" had previously sought to escape, but some, like the leaders of IFM, were in trouble with the regime for their political opposition. The exchange rate for prisoners averaged at least 1,000 DM

per person throughout the 1970s, with the ransom price gradually increasing. For 1977, Amnesty International (1978) reported a range of 30,000–160,000 DM per prisoner. Cost estimates for the FRG range up to a total of 1 billion marks for the entire period (Naimark 1992, p. 79). Over 30,000 persons were "bought free" between 1961 and 1989 (*Frankfurter Allgemeiner Zeitung*, Oct. 21, 1989).

The opening of opportunities in the Soviet Union and increasing dissatisfaction in the GDR led to a marked increase in both the number of applications and illegal departures by the late 1980s. Despite a substantial increase in the number of permits granted, there was a new desperation among those who sought to leave by 1989. In May, Hungary set the stage for mass exodus by removing fortifications along its border with Austria. With a new avenue of escape, the number of migrants from East to West doubled from 5,887 in April to over 10,000 in July. The number of departures continued to accelerate throughout the summer until August when they doubled again with returning vacationers.

By the end of September 1989, a mass exodus of over 100,000 persons had exceeded all previous years. The embarrassment to the GDR was made all the worse by the impending celebration of its fortieth anniversary to be attended by Mikhail Gorbachev in early October. Practical considerations could not be ignored, as some major service areas became seriously handicapped by the loss of workers. Under the circumstances, Honecker was persuaded to accept a compromise in secret negotiations at the United Nations (Jarausch 1994, p. 21). This led to a series of half steps and face-saving measures by the regime until the border was finally demilitarized in November. Almost 350,000 East Germans left for the FRG in 1989, followed by another 240,000 from January to June 1990, when reunification was assured (Mehrlander 1994, p. 6).

This account of emigration's role in opposition to the GDR regime is one-sided in its focus on those who left or sought to leave either legally or illegally for a democratic, capitalist, yet German society where they were guaranteed immediate citizenship. Such an emphasis reflects the fact that exit was the major expression of mass dissatisfaction throughout the short history of the regime and played a decisive role in precipitating the autumn crisis of 1989. Against this background, I identify critical changes in the exit repertoire that indicate the way in which the repressive nature of an authoritarian regime leads to the reappropriation of earlier, traditional forms of action. To this end, the massive movements of tens of thousands of families, neighbors, and kin are disaggregated into specific exit episodes over the period 1961–89. They illustrate the increasing incorporation of new elements as the opportunities for exercising the modern repertoire expanded and the cohesion of the Warsaw Pact countries disintegrated.

HYBRID REPERTOIRES OF EXIT

Examination of collective actions that were originally clandestine requires relaxation of most strictures of event analysis methodology that provide guidelines for obtaining data in more open societies (Tilly 1986, 1995b; Tarrow 1989). Recommended procedures for population enumeration from newspapers checked against alternative data sources, systematic sampling of years, months, or days, and standardized coding schemes (Rucht, Koopmans, and Neidhardt 1998; McCarthy, McPhail, and Smith 1996; Olzak 1989) fail the test of practicality for analyzing exit actions within Leninist regimes. Instead, I rely on monographs, the accounts of journalists, and the reports of human rights organizations for information on exit episodes that occurred before 1989. For 1989, six international daily newspapers from West Germany, Great Britain, and the United States were systematically coded for all protest events (for a description of this data set, see Mueller [1997]). Only those events that reflect the exit repertoire are included here. The newspapers are the *Suddeutsch Zeitung* (*SZ*), the *Frankfurter Allgemeiner* (*FAZ*), the *Times* of London (*LT*), the *Guardian* (*G*), the *New York Times* (*NYT*) and the *Washington Post* (*WP*). Given the unsystematic nature of the early sources, there is no attempt to measure the prevalence or change over time of specific combinations of repertoire elements.

Exit actions can be analyzed in terms of the same elements that have been identified earlier to characterize profiles of collective action marking the transition from a traditional to a modern repertoire. Throughout the period under consideration, the elements are reconfigured to form hybrid patterns of action that change from a preponderance of traditional to primarily modern repertoire elements depending on two sets of factors: (1) changes in political opportunities within the GDR due to shifts in relationships with allies in the Warsaw Pact and the Soviet Union as well as the broader international community, particularly West Germany, and (2) the evolving action repertoires of groups using "voice" to express their grievances. The initiative for testing political constraints shifted back and forth between those who sought to leave (*Ausreiser*) and reformers in challenging the state with new hybrid combinations.

Exit Hybrid 1: Clandestine Escapes

Escape attempts began as soon as Berliners discovered the border was being sealed in the early hours of August 13, 1961 (Kemp 1987). In the first months, individuals, couples, and families could crawl under barbed wire, jump from windows in adjoining buildings, climb through the sewers, or swim the Tetlow Canal. As fortifications along the border were secured and the number of border guards increased, escape became more

difficult. Although individual efforts continued until the fall of 1989, within a year, it was increasingly necessary to rely on the expertise of Westerners—both idealistic "escape helpers" and criminal syndicates devoted to the same purpose for sometimes exorbitant profit. In the early 1960s, there were several thousand helpers from the West working voluntarily to assist those who sought to escape. One of the most famous of these was an idealistic young student by the name of Victor Fuchs, known as "Tunnel Fuchs" for the seven tunnels whose construction he organized from West to East Berlin. The most famous was Tunnel 57.

> *October 3–5, 1965* "Tunnel 57" was the last and the most famous of the tunnels constructed by Victor Fuchs and his many helpers, most of them students from either the Free University in West Berlin or from outside the city. The name was derived from the 57 individuals who were brought through the 400 foot tunnel 20 feet under the surface of the city on the nights of October 3 through 5, 1965. It had taken 40 individuals six months to construct a tunnel that originated in a deserted bakery on Bernauerstrasse in West Berlin and extended into an empty block of apartment buildings on the east side of the wall. Financing for the elaborate construction—with electric lighting, artificial ventilation and timber supports for the water-soaked walls—was provided by a West German magazine and by production of a film detailing the construction and escapes. After 57 had escaped, the tunnel was discovered and destroyed by police at the east entrance. (Paraphrased from accounts given in Kemp [1987, pp. 67–82], and Galante [1965, pp. 256–59].)

Although many teams worked to help East Germans in the early years after 1961, none were so successful as Fuchs. These early escape projects received either the tacit or overt support of the West German police, a level of cooperation that was gradually withdrawn in the late 1960s as the FRG made regularization of relations with its eastern neighbor a top priority (Ash 1993, p. 142; Kemp 1987, pp. 77–78).[16]

Tunnel 57 captures the hybrid nature of clandestine escapes in the first few years after the wall was constructed, when the number of escapes over the East-West border was at its peak (*Frankfurter Allgemeiner Zeitung,* October 21, 1989). The requirement of secrecy was the key factor influencing the clandestine pattern of this profile (see table 1, col. 2). Secrecy influenced other key requirements: actors were limited to those who

---

[16] As risks increased, escapes were increasingly in the hands of "professionals" who combined a taste for adventure with technical skills in forging false passports and diplomatic papers, devising ingenious disguises, and modifying a wide range of vehicles to hide passengers. Although a few mercenaries made small fortunes in this business, the "helpers" reflected a range of motives and continued to include idealists who charged only enough to cover the expenses of escape. These teams of specialists provide an odd twist to McCarthy and Zald's emphasis on the professionalization of social movement activists (1973)

could be trusted based on the intimate, personal knowledge of family, friends, and neighbors; the division of labor among those who escaped originated in their primary relationships; and action was direct in terms of securing the goal through the activity itself. In the West, requirements for secrecy were much less decisive. The several thousand "Victor Fuchs" in the West shared a common purpose that was both humanistic and political, but they seldom knew one another. Fuchs was a self-appointed activist-entrepreneur whose multiple successes in escape efforts established a reputation sufficient to attract large capital investments. The unusual media attention he received insured that the drama of the escapes was brought to a larger audience. Reception of West German television in most of the GDR encouraged others to believe that escape was possible. Yet, the repertoire is primarily traditional rather than modern. Based in primary relations, actions are physical and direct in the achievement of collective goals.

### Exit Hybrid 2: The Helsinki Accord Petitioners

High levels of repression characterized the early 1970s in the wake of the 1968 Warsaw Pact invasion of Czechoslovakia that ended the reform attempts of Prague Spring for "socialism with a human face." Independent political activity came to a standstill. By the middle of the decade, however, the *Ausreiser* had become increasingly assertive due to both general and specific sources of encouragement.[17] The general relaxation of East-West tensions was reflected in superpower détente, the *ostpolitik* of Willy Brandt's social-liberal coalition, and the signing of the UN Basic Treaty, which recognized both German states and granted the GDR long-sought

---

[17] These various "sources of encouragement" fit only the newly defined dimensions of international political opportunity rather than the more typical features of domestic political systems discussed by political process theorists (see, esp., McAdam [1996] and Oberschall [1996] in McAdam, McCarthy, and Zald [1996]). The imperial needs of the Soviet Union for international recognition of its satellites in Eastern Europe and the continuing economic crises faced by all of the Warsaw Pact countries during the Brezhnev years could not be characterized as reflecting instability in GDR elite alignments, the presence of domestic elite allies for *Ausreiser* or opposition activists, or a change in the state's propensity for repression (McAdam 1996). Also, the "relative openness or closure of the *institutionalized* political system" (McAdam 1996, p. 27) of the GDR changed little when Gorbachev began dismantling much of the machinery of the Communist Party of the Soviet Union (CPSU) and like-minded reformers followed suit in a few selected countries like Hungary. The modest support provided by the West, such as investigations by committees of the U.S. Congress and human rights organizations, could be interpreted as representing the work of an external configuration of allies, but "political opportunity structure" usually characterizes the political system within which collective action occurs.

international recognition (Maier 1997, pp. 25–26). More specifically, *Aus-reiser* were emboldened in the mid-1970s by the ongoing negotiations and human rights provisions of the Helsinki Final Act that was signed by the Soviet Union and other Warsaw Pact countries (Thomas 1997). For the West, these moves sought to stabilize Cold War relations and "to improve the lot of ordinary citizens in Eastern Europe by dealing with their single party regimes" (Maier 1997, p. 94). For the East, they reflected concessions that would legitimate the postwar boundaries in Eastern Europe and secure vital technology and credits from the West.[18]

The Helsinki Accords were signed in August 1976, and the next 18 months saw an unprecedented and unexpected outpouring of collective action in the Soviet Union and much of Eastern Europe centering around the petition process and culminating in the signing and publication of Charter 77 in Czechoslovakia, which became the symbol for a new era in human rights activism in the Soviet bloc (Thomas 1997, p. 119).[19] Although the GDR saw very little of the activity that resulted in the Helsinki

---

[18] By the late 1960s, both the Soviet Union and the Leninist regimes of Eastern Europe faced the increasing calcification of their economies following the Warsaw Pact suppression of Prague Spring that brought an end to the reform efforts of the decade. The agreements signed in the early and mid-1970s would not only sanctify the postwar boundaries but would also bring in the advanced technology to help make East European goods competitive in international markets as they sought hard currency for subsidizing the Western consumer goods needed to satisfy increasingly restless populations (Chirot 1991). Maier's research in the recently opened archives of the GDR indicates just how desperate the situation had become by 1989 when Gerhard Schurer, head of the Planning Commission, was finally permitted to report to Egon Krenz, Honecker's successor, that the debt service alone on foreign loans was over 60% of yearly export earnings (Maier 1997, p. 60). Schurer traced this financial debacle to the Eighth Party Congress, in 1971, which resolved that the GDR must support a generous consumer society and welfare state to avoid unfavorable comparisons with West Germany. As foreign indebtedness began to mount in the 1970s, Honecker refused to permit Schurer to report.

[19] In the first all-European meetings since the end of World War II, the 35 member states of the Commission on Security and Cooperation in Europe (CSCE) had acceded to revolutionary proposals by the smaller states of the European Community for a definitive commitment to human rights and the self-determination of people (Thomas 1997). In exchange for the agreement on the inviolability of frontiers and nonintervention in their internal affairs as well as extensive provisions for scientific and technological exchanges, the Soviets and East European states grudgingly signed international protocols accepting a standard for judging their own behavior outside the ideological superstructure of Marxism-Leninism (Schopflin 1993, p. 179). As a safeguard on implementation of the agreement, the signatory states would reconvene at intervals for evaluations. Proceeding *as if* the official commitments were sincere, dissidents in Poland, Czechoslovakia, and the Soviet Union petitioned their governments for the release of political prisoners and other denials of human rights protected in the Helsinki Final Act. With 242 original signatures, Charter 77 was a four-page document chronicling human rights abuses in Czechoslovakia and calling for dialogue with authorities (Thomas 1997, p. 202).

Watch committees in the Soviet Union, Poland, and Czechoslovakia, it was in this political climate that the Riesa petition was circulated, signed, and publicized.

> *1976. The Riesa Petition.* Thirty-three people from Riesa, near Dresden in Saxony, tested the limits of the new international agreement by submitting a petition to the United Nations, to the signatory nations of the Helsinki agreement, and to "world public opinion" seeking to renounce their East German citizenship and move to West Germany that they might "secure full human rights, including that of free choice of domicile and place of work." Another forty-six East Germans added their names in the next few weeks. Although the leader of the petition, Dr. Karl-Heinz Nitschke, was first demoted at work and then arrested and detained for a year, all were eventually urged to withdraw their applications before they were finally released in the west where Nitschke had disclosed the petition to the media. (Paraphrased from accounts in Thomas [1997, pp. 139–40] and Woods [1986, p. 34, petition on p. 178].)

After the modest success in recruiting additional signatures to the Riesa petition, there is no evidence of further public petitions for exit until 1984 when eight *Ausreiser* sought asylum at the U.S. Embassy in East Berlin with a signed statement addressed to President Reagan threatening a hunger strike as "our last desperate attempt to make known our desire to emigrate to a democratic country" (Woods 1986, pp. 172–73).

The Riesa petition marked a major break with escape attempts and prisoner exchanges of the pre-Helsinki period, although both continued unabated. With public petitions, exit claims emerged in the open arena of political struggle. The petition was published as a public appeal to international third parties to pressure the GDR under the blanket authority of the Helsinki agreements and to fellow *Ausreiser* to join in publicizing the petition strategy for leaving the GDR. The petition at Riesa was conspicuous as an open claim for exit, but given its dramatic though limited success as a modular strategy, the larger effect of the agreement is visible in the enormous increase in personal petitions for exit visas accompanied by appeals to the Helsinki provisions for freedom of movement (Allen 1991, pp. 133–34). The number of applications rose to 200,000 in 1977 and 300,000 by 1978 (Thomas 1997, p. 215). The Honecker regime once again demonstrated its lack of sensitivity to the international pressures that had led to Helsinki by refusing to increase the number of permissions granted.

Exit Hybrid 3: The Open Vigil

For the next change in the exit repertoire, the autonomous (without state sponsorship) peace movement modeled new initiatives in testing the limits

of collective action. Influenced by the new social movements in Western Europe, and particularly the FRG, peace activists dramatized their grievances with petitions that had become "modular" (Tarrow 1993) throughout the Eastern bloc during the Helsinki-inspired human rights mobilization. The Dresden youths' Social Service for Peace declaration attracted 6,000 signatures and called for alternatives to military service in May 1981. The highly controversial Berlin Appeal the following January by Robert Havemann, long-time leader of GDR dissidents, and the popular East Berlin minister, Rainer Eppelmann, was eventually signed by over 2,000 persons in the GDR and West Europe. The large number of signatures marked a brief period of more open defiance of the state.

The next step was to move from the signed petitions to demonstrations in public forums attached to official events. The most dramatic of these was the Dresden Forum and alternative march in February 1982 by 5,000 young people who gathered under the protective canopy of the GDR's annual commemoration of the losses from Allied bombing to call for alternatives to military service. Following events in Dresden there was a brief flurry of autonomous peace actions involving thousands of young people in many cities. Women soon mobilized as well after the state promulgated a new military service law authorizing their conscription for unarmed military service (Allen 1991, pp. 93–124).

Informed by their contacts with the West European peace movement, women and young people adapted the petition, the march, and the open meeting associated with official events that were guaranteed coverage by Western media. A major center of these actions was Jena, a university town with the reputation as a center of cultural activity and critical thinking outside the framework of church or party. In December 1982, peace activists planned to stand in the central square for a moment of silence on Christmas eve. Police intervention resulted in the arrest of 14 leaders, followed immediately by an indignant public response in Western Europe that prompted their release. This success was short-lived, however. After leaders sent a letter to Honecker protesting local police actions, 24 of the active core of the Jena peace movement were expelled to West Germany in May and June, pointedly demonstrating the limits of official tolerance for activity outside church auspices (Allen 1991, pp. 113–14). Nevertheless, *Ausreiser* in Jena now embraced the forms of action developed by the peace movement since emigration was, in fact, their goal.

> *1983. The Jena Vigils.* For eight weeks, silent vigils were held each Saturday outside the government offices that processed applications for permission to leave the GDR. Applicants and their supporters wore white, becoming known as the "White Circle." Initially, there were 30 *Ausreiser,* but their numbers soon swelled to 200. After their arrest and imprisonment, Amnesty

International received reports of *Ausreiser* demonstrating in other towns where participants were also arrested. (Paraphrased from accounts in Amnesty International [1984], Woods [1986, p. 34] and Torpey [1995, p. 72].)

Like the Riesa petition, the Jena White Circle showed both innovation and adaptation of a basic form of action, in this case, symbolic meanings attached to the public demonstration. "White" as well as the immobile vigil form symbolized purity and nonviolence in contrast to state repression; "silence," the absence of a public forum for open discussion; and "the circle," group solidarity in the face of legal restrictions on freedom of movement, of intransigent bureaucracies, and of state repression. Also like the *Ausreiser* in Riesa, they sought to attract the attention of Western media, thereby creating sufficient international pressure to persuade the regime to permit some of the demonstrators to leave the country. The success of their venture was realized in their expulsion as well as the spread of similar demonstrations to other cities.

These two episodes by *Ausreiser* exhibit many elements of the modern repertoire in a state without a functioning electoral system. The interaction with authorities in such a system occurs not through national level demonstrations that appeal to coalitions of voters who might influence the state directly, but to an international audience of interested states and nongovernmental organizations (NGOs) that bring pressure to bear indirectly on regime authorities by providing or withholding valuable resources such as trade, hard currency, or the relaxation of hostilities. Although the petition in Riesa and the vigils in Jena were successful in attracting local supporters beyond the original core of activists, state control of communication and association prevented these events from leading to the formation of social movement organizations or coordinated campaigns to pursue exit goals on a broader scale. The level of generality or modularity was still limited by a high level of repression and by the transient nature of the *Ausreiser*'s participation. In addition, goals were confined to the personal preferences for exit of a relatively small number of people. There were still no attempts by *Ausreiser* to bring about changes in state regulations limiting travel and emigration, although protests were increasingly framed in terms of broad human rights promulgated in international accords signed by the GDR.

In the six years from 1983 until 1989, oppositional activity in the GDR repeated again and yet again, the pattern set by the autonomous peace movement of laying a foundation for mobilization within sympathetic local churches, of extending the forms of action as well as critical analysis beyond levels acceptable to the hierarchies of church and state, and then of retrenchment after heavy state repression. As this pattern was repeated on issues of ecology and then human rights, activists were gradually radicalized. Groups long active on these issues in East Berlin joined together

in creating a formal organization, the Initiative for Peace and Freedom, the first oppositional organization clearly outside the protection of the church. Influenced by their own continuing experiences of repression, by the activist models of Charter 77 in Czechoslovakia and the Greens in West Germany, and by Gorbachev's *glasnost* efforts in the Soviet Union, activists in the GDR had finally reached the point of invoking "voice" to call for democratization of their socialist society.

Exit Hybrid 4: Open Demonstrations and Use of the "Rights" Frame

Although repression never declined for long, cries of outrage from within the GDR and from the international community accelerated at each new episode of arrest, detainment, and expulsion. Two waves of repression, one in November 1987 against new *samizdat* publications and again in January 1988 against counterdemonstrators at the annual Rosa Luxemburg remembrance, were met by protest meetings, vigils, prayer services, and other solidarity actions in 32 cities and towns across the GDR with tens of thousands of participants (Allen 1991, pp. 169–71). Hundreds of members of the democratic opposition across the Soviet bloc signed a joint declaration addressed to the leaders of the GDR condemning the arrests and expulsions of initiative leaders who led the counterdemonstration. Among those arrested at the Rosa Luxemburg demonstrations were members of an East Berlin *Ausreiser* group who were also expelled (Torpey 1995). As the number of public protests gradually accelerated in early 1989, the *Ausreiser* played an increasingly prominent role. At this point, demands for personal exit had been redefined in terms of universal rights to freedom of movement. Although exit actions increasingly resembled the contemporary repertoire, the human rights frame was invoked for audiences that could only indirectly bring pressure to bear on the leadership of the GDR.

> *March 1989. Demonstrations at the Leipzig Trade Fair.* Twice a year, the Western media visited Leipzig for its semiannual Trade Fair and provided an opportunity for local dissidents to present their claims to Western audiences. On the first night of the Fair, several hundred demonstrators carried posters for travel rights and were arrested and roughed up before Western television cameras. Again on March 13, the Trade Fair attracted over 650 demonstrators who were attacked by the Stasi as they chanted, "Freedom" and "Human rights." Security forces reported that over half of the demonstrators were would-be emigrants, carrying posters with messages such as "Travel Freedom-Not Official Arbitrariness." (Paraphrased from accounts in *SZ, FZ, WP*, Maier [1997, p. 135], and Torpey [1995, p. 150].)

The mass exodus of 100,000 East Germans through Hungary and Czechoslovakia at the end of the summer emboldened *Ausreiser* still within the GDR to demand the *right* to travel or to emigrate as citizens of the GDR.

> *September, 1989. Demonstrations at the Leipzig Trade Fair.* On September 4, as the Nicolai Church's Monday night peace prayer meetings resumed in the middle of the Trade Fair, *Ausreiser* clashed with newly galvanized reformers in competing chants of "We want out!" and "We are staying." The church was surrounded by police, once again playing their historic role before West German television cameras of arresting and beating demonstrators leaving the church. (Paraphrased from accounts in *SZ, FZ, LT,* and Maier [1997, p. 138].)

On the eve of the GDR collapse, public demonstrations in the face of police brutality combined with grievances framed in the universal language of human rights have most of the elements of the modern repertoire. The absence of formal organizations was due to the constant attrition of members as *Ausreiser* left the GDR, although some made up the fringes of reform groups where they were tolerated (Torpey 1992).

Exit Hybrid 5: Mass Exodus and Modifications of the Modern Occupation

By early fall, protest demonstrations within the GDR assumed a pattern of action that varied little except in the level of police violence from the modern repertoire; and police violence effectively ended with an impromptu settlement in Leipzig on October 9. The right to travel or emigrate remained one of the major grievances expressed in the fall demonstrations until the Berlin Wall was opened one month later. Outside the GDR, attempts to emigrate to the West took on the versatility and innovation associated with a highly mobilized cycle of protest (Tarrow 1993).

Thousands of East German vacationers returning in the late summer took advantage of the Hungarian border opening to make impromptu changes in their final destinations. By the end of August, 150,000 East Germans were in Hungary trying to escape illegally (Maier 1997, p. 126). When the border to Hungary was then closed off by the GDR, *Ausreiser* adapted new forms of exit actions (Oberschall 1996). For the first time, the type of actions benefited from the mass numbers attempting to leave the GDR without permission. The first and most creative innovation was a variation on the factory occupation or student sit-in.

With Czechoslovakia as the one remaining ally of the GDR with whom there was still an open border for East Germans, thousands of young families and groups of friends made their way to Prague.

> *Late September-Early October, 1989. Occupation of the German Embassy in Prague.* For eleven weeks, *Ausreiser* families and single adults sought asylum and free passage to West Germany by occupying the building and grounds of its embassy in Prague where they received diplomatic immunity. By the first weekend in October, the numbers had swelled to over 5,000. The presence of a significant number of children in both the embassy and

nearby shelters created a "humanitarian crisis" due to the lack of food, shelter and sanitary facilities. (Paraphrased from accounts in *SZ, FAZ, LT, G, NYT, WP.*)

Following the purpose of the modern occupation in creating a crisis for the targets of action, widespread media coverage of this embarrassing scene on the eve of the GDR's fortieth anniversary celebration led to a UN-negotiated compromise whereby the Honecker regime would "expel" those occupying the embassy, and send them by train back through the GDR and out to West Germany branded as "undesirables." This attempt to diffuse the situation was only momentarily successful. GDR permission for special trains to carry *Ausreiser* to West Germany led to a flood of new arrivals. Telecasts and new pictures were beamed around the world of East Germans scaling embassy walls, pulled up by West German personnel on the inside, with Prague police trying to pull them back down. The presence of pregnant women and small children added to the drama. Enormous numbers created another humanitarian crisis, and by the middle of the next week 11,000 *Ausreiser* had been carried by sealed trains back across the GDR to resettlement camps in West Germany.

Exit Hybrid 6: Mass Numbers and the Urban Insurrection

With the fortieth anniversary of GDR founding and the arrival of Gorbachev imminent, Honecker closed the border to Czechoslovakia, thus creating a sense of desperation on the part of *Ausreiser* stranded now in Dresden, the GDR's major gateway by rail to Czechoslovakia and its West German embassies. The next modification of the modern repertoire was an adaptation of the urban insurrection.

> *October 4–9, 1989. Riots at the Dresden Train Station.* By October 4, the Dresden train station had been emptied and barricaded to keep out *Ausreiser* who might try to board the eight trains carrying East Germans who had occupied the West German embassies in Prague. Police used billy clubs, water cannons, and tear gas on the crowd of 10,000 who took to the streets and hurled back cobblestones as they barricaded the mayor in his office. In the ensuing riot, one protester was killed and dozens were injured and arrested. The next day, the crowd had grown to 30,000. By the eighth, it appeared that violence would become more lethal until word spread that the mayor had received delegations of churchmen and citizens, would release those arrested, and begin conversations with the community, the first party leader to respond to the growing demand for "dialogue." With a spirit of celebration, over 22,000 filled four churches in the center of Dresden on the night of October 9. (Paraphrased from accounts in *FAZ, WP, NYT,* Maier [1997, p. 145], Jarausch [1994, p. 22], and Naimark [1992, p. 89].)

Street violence, barricades, growing participation—East Germany had not seen such open rebellion since 1953, but this time there was not the

possibility of Soviet intervention. In Leipzig, the citizens were also willing to risk what they had not dared before, and 70,000 braved the threat of a "Tiananmen Square" solution for a demonstration that marked the end of police violence as a response to peaceful gatherings in the GDR. As the police backed down, both the size and number of demonstrations escalated and spread throughout the country. By October 18, Honecker and several members of the Politburo had resigned.

Due to the exit crisis, travel and emigration remained the top issues confronting the regime until November 9. On November 4, close to 1 million men, women, and children came together in Alexanderplatz, or "Alex," the heart of East Berlin, called together by the country's leading writers, artists, and intellectuals for the largest demonstration of the fall. Sharing the platform with leaders of both New Forum, a recently formed civic organization, and the SED, above all else speakers called for the end of travel restrictions. Two nights later, in Leipzig, mass demonstrations presciently called not only for the end of travel restrictions, but also for the end of the Berlin Wall, as well as the constitutional role of the party (Maier 1997, p. 159). The next day, both the Central Committee of the SED and the Politburo resigned after an unprecedented rebuff of their proposal for travel legislation by the *Volkskammer*, the hitherto rubber-stamp legislative branch of government. As the new leadership of the party tried one halfway measure after another, reopening the border to Czechoslovakia resulted in yet another occupation of the West German embassy and a diplomatic stalemate between the two die-hard Leninist regimes as tens of thousands of East Germans once again surged across the border. In the confusion of the next two days, increasingly uncertain party leaders inadvertently opened the Berlin Wall and the frontier between the two Germanys.

CONCLUSION

The focus here has been atypical for recent work on the GDR. Despite the critical importance of flight or escape in the 1989 fall of the GDR's Leninist regime, an explanation of those momentous events is not the central concern here. Rather, the theoretical task has been to expand theories of contentious repertoires (Tilly, Tarrow, and others) to fit the repressive conditions of authoritarian regimes by examining collective escape episodes as hybrid forms of collective action. The method has been to disaggregate key dimensions of traditional and modern repertoires of action and then to identify their multiple hybrid forms in the GDR over almost 30 years of varying constraints and opportunities.

Over the period 1961–89, collective actors pursuing exit and voice in the GDR alternated in their experimentation with new hybrid forms of

action. Reformers sometimes showed the way—with petitions after Helsinki and with vigils during the peace movement. However, it was the *Ausreiser* who innovated the tunnels of the 1960s, the open display of chants and banners in demonstrations associated with the Leipzig trade fairs of 1989, the embassy occupations in the summer, the rediscovery of the urban insurrection at Dresden in the fall, and, finally, the "happening" at the Berlin Wall on November 9. Over these 28 years, private claims for escape were transformed into the public and universal language of the right to travel and to emigrate.

In reviewing a series of collective efforts to leave a Leninist regime, I have identified key elements for defining profiles of collective action repertoires as criteria for evaluating whether these efforts fall within the domain of political process theory. Based on such criteria drawn from the work of Charles Tilly and Sidney Tarrow contrasting traditional and modern "collective action repertoires" or "repertoires of contentious politics," this particular cluster of exit episodes represent variable hybrid combinations of both repertoires due to the varying constraints on action of the GDR's Leninist regime. Such authoritarian regimes not only monopolize power, authority, and the right to make claims against the state and third parties, but they also seek to preempt the means of communication and association on which the modern repertoire is based. In such regimes, Hirschman (1986, 1993) came to realize that the pursuit of exit is a likely form of collective action and as hazardous as that of "voice."

By disaggregating and recombining the defining elements of modern and traditional repertoires to include collective actions such as exit, it is possible to expand the domain of political process theory to cover a broad range of nondemocratic regimes in which civil society has been severely restricted. In any authoritarian regime, the hybrid form of these combinations will be variable at any given time and place, depending on constraints and opportunities, resources, collective identities, and the previous history of struggle.

Although it is beyond the scope of this article to specify the configuration of constraints and opportunities that make any of the multiple forms of exit attractive in authoritarian regimes, critical factors in other societies can be noted whose ethnic communities have either been divided or have fled into expatriate enclaves as a result of revolution, warfare, and military occupation. Restricting consideration to Leninist regimes, the partition of Vietnam and Korea stand out as examples that most closely parallel the division of post–World War II Germany. Ethnic enclaves of Tibetans in India and Cubans in the United States reflect total loss of the homeland due to either foreign occupation or revolution. In all four cases, an established Leninist regime placed severe but varying levels of constraint on cross-border travel as well as on emigration.

These constraints have given rise to multiple exit profiles, ranging from the dramatic flight of the Dalai Lama and his government into exile (Avedon 1984) to the mass exodus of Cubans to Florida (Aguirre 1995, 1994). Regardless of later hybrid forms, the initial flight in both cases drew on the traditional repertoire, acting directly on grievances through physical action supported by trusted intimates. Among divided societies, North Korea remains one of the most isolated Leninist regimes of the 20th century, with such tightly controlled borders that those few escaping to China, Russia, Japan, or South Korea do not enter the ledger of world refugee statistics (United States Committee for Refugees 1999). To the extent that escape has been possible, over time the collective action profiles have tended to include a variety of hybrid elements. Flight from other types of authoritarian regimes would probably reveal even greater variety.

In the case of the GDR, the episodes cited here demonstrated a transition from more to less traditional elements. Is this also a progressive transition? Tilly argues in his book on contention in Britain (1995b) that the transition from a traditional to a modern repertoire was not an "unmitigated advance" (p. 376). Historically, in Europe and North America, it sacrificed short-term visible results based on the immediate moral pressure of small-scale communities to the long-term, indirect benefits of those whose interests could be advanced through the larger scale of specialized associations and social movements operating at a national level. Yet, such a sacrifice is not obvious in contemporary authoritarian regimes, where repression may have localized social relations in traditional social niches but the pervasiveness of state surveillance makes it almost impossible for work groups, neighborhoods, or housing estates to act openly on the basis of shared grievances. In such regimes, the modern repertoire or critical elements of it have a strategic appeal.

As political opportunities appeared in the GDR due to external shifts in the international framework of Warsaw Pact states within which the GDR regime existed, *Ausreiser* as well as reformers experimented with ever more modern forms of action. Particularly, they used the media to engage the support of NGOs and human rights activists who would invoke a variety of international pressure points to influence the Honecker regime. As long as the repressive apparatus of the regime was in place, major solutions could only occur at the border or through neighboring Leninist countries, and there with foreign assistance. Without these elements of indirect influence that characterize the modern repertoire, the major changes achieved through the mass exodus of 1989 could not have been achieved.

The GDR's mass exodus of 1989 led one of the most perceptive Western observers of Eastern Europe's velvet revolution to characterize the GDR's "emigration movement" as analogous to Solidarity in Poland (Ash

1993, pp. 195, 531). For students of contentious repertoires, Ash's characterization appears at first to ignore critical differences in the forms of collective action. Yet, reflection suggests that both Solidarity, a labor union/political movement, and emigration, a collective and episodic pursuit of "exit," functioned similarly as vehicles for achieving the most deeply felt grievances of the mobilized masses.

REFERENCES

Aguirre, Ben E. 1995 "Culture of Opposition and Collective Action in Cuba " Paper presented at the annual meetings of the American Sociological Association.
————. 1994. "Cuban Mass Migration and the Social Construction of Deviants." *Bulletin of Latin American Research* 13:155–83.
Avedon, John F. 1984. *In Exile from the Land of Snows.* New York Alfred A. Knopf.
Allen, Bruce. 1991. *Germany East: Dissent and Opposition.* Montreal: Black Rose Books.
Amnesty International. 1974–89. "The GDR " *Annual Report.* London Amnesty International.
Arendt, Hannah. 1951. *The Origins of Totalitarianism* New York: Harcourt, Brace.
Ash, Timothy Garton. 1990. *The Uses of Adversity* New York: Random House, Vintage Books.
————. 1993. *In Europe's Name: Germany and the Divided Continent.* New York: Random House, Vintage Books.
Bach, R. L., and L. A. Schram 1982. "Migration, Crisis and Theoretical Conflict." *International Migration Review.* 16:320–41.
Barry, Brian. 1974. "Review Article Exit, Voice, and Loyalty." *British Journal of Political Science* 4·79–107
Boudreau, Vince. 1996. "Northern Theory, Southern Protest: Opportunity Structure Analysis in a Comparative Perspective " *Mobilization: An International Journal* 1: 175–89.
Brockett, Charles D. 1991. "The Structure of Political Opportunities and Peasant Mobilization in Central America," *Comparative Politics* 23.253–74.
————. 1995. "A Protest-Cycle Resolution of the Repression/Popular-Protest Paradox." Pp. 117–44 in *Repertoires and Cycles of Collective Action,* edited by Mark Traugott. Durham, N.C.: Duke University Press.
Childs, David. 1988 *The GDR: Moscow's German Ally,* 2d ed. London: Unwin Hyman.
Chirot, Daniel. 1986. *Social Change in the Modern Era.* New York: Harcourt, Brace, Jovanovich.
————. 1991. "What Happened in Eastern Europe in 1989?" Pp. 3–32 in *Leninism and the Decline of the Left,* edited by Daniel Chirot. Seattle: University of Washington.
Dallin, Alexander, and George W. Breslauer 1970 *Political Terror in Communist Systems.* Stanford, Calif.: Stanford University Press.
Darnton, Robert. 1991 *Berlin Journal, 1989–1990.* New York: W.W. Norton.
Di Caprio, Lisa. 1990. "East German Feminists: The Lila Manifesto." *Feminist Studies* 16·621–34
Finer, Samuel E. 1974. "State-Building, State Boundaries and Border Control in Light of the Rokkan-Hirschman Model." *Social Science Information* 13:79–126.
Galante, Pierre 1965. *The Berlin Wall.* Garden City, N.Y.: Doubleday.
Gamson, Bill. 1992. "The Social Psychology of Collective Action." Pp 53–76 in *Frontiers in Social Movement Theory,* edited by Aldon Morris and Carol M Mueller. New Haven, Conn.: Yale University Press

Hankiss, Elmer. 1990. "What the Hungarians Saw First." Pp. 13–36 in *Spring in Winter: The 1989 Revolutions,* edited by Gwyn Prins. Manchester: Manchester University Press.

Hirschman, Albert O. 1970. *Exit, Voice and Loyalty.* Cambridge, Mass.: Harvard University Press.

———. 1986. "Exit and Voice: An Expanding Sphere of Influence." Pp. 77–101 in *Rival Views of Market Society and Other Recent Essays* by Albert O. Hirschman. New York: Penguin Books

———. 1993. "Exit, Voice, and the Fate of the German Democratic Republic: An Essay in Conceptual History." *World Politics* 45 173–202.

Hirszowicz, Maria 1986. *Coercion and Control in Communist Society.* Brighton: Wheatsheaf Books

Holmes, Leslie. 1986. *Politics in the Communist World.* Oxford: Clarenden Press.

Jarausch, Konrad H. 1994. *The Rush to German Unity.* New York: Oxford University Press.

Johnston, Hank. 1991. *Tales of Nationalism: Catalonia, 1939–1979.* New Brunswick, N.J.: Rutgers University Press

Joppke, Christian. 1995. *East German Dissidents and the Revolution of 1989: Social Movements in a Leninist Regime.* New York: New York University Press.

Jowitt, Ken. 1978. "Inclusion and Mobilization in European Leninist Regimes." *World Politics* 28:69–97.

———. 1992. *New World Disorder.* Berkeley: University of California Press.

Kemp, Anthony. 1987. *Escape from Berlin.* London: Boxtree.

Klandermans, Bert. 1984. "Mobilization and Participation: Social Psychological Expansions of Resource Mobilization Theory." *American Sociological Review* 49: 583–600.

——— 1997. *The Social Psychology of Protest.* Cambridge, Mass · Blackwell.

Lane, David. 1984. "Human Rights under Socialism." *Political Studies* 32:34–68.

Laponce, Jean. 1974. "Hirschman's Voice and Exit Model as Spatial Archetype." *Social Science Information* 13:67–81.

Lemke, Christiane. 1986. "New Issues in the Politics of the German Democratic Republic: A Question of Political Culture?" *Journal of Communist Studies* 2:341–58.

Lichbach. Marc 1995. *The Rebel's Dilemma* Ann Arbor: University of Michigan Press.

Lipsky, Michael. 1968. "Protest as a Political Resource." *American Political Science Review* 62:1144–58.

Maier, Charles S. 1997. *Dissolution: The Crisis of Communism and the End of East of German.* Princeton, N.J.: Princeton University Press.

McAdam, Doug. 1982. *The Political Process and the Development of Black Insurgency.* Chicago: University of Chicago Press.

———. 1983. "Tactical Innovation and the Pace of Insurgency." *American Sociological Review* 48·735–54.

———. 1996. "Political Opportunities: Conceptual Origins, Current Problems, Future Directions." Pp. 23–40 in *Comparative Perspectives on Social Movements,* edited by Doug McAdam, John McCarthy, and Mayer Zald. New York· Cambridge University Press.

McAdam, Doug, John D. McCarthy, and Mayer N. Zald. 1996. *Comparative Perspectives on Social Movements.* New York: Cambridge University Press.

McAdam, Doug, Sidney Tarrow, and Charles Tilly 1996. "To Map Contentious Politics." *Mobilization: An International Journal* 1:17–34.

McCarthy, John D., Clark McPhail and Jackie Smith. 1996. "Images of Protest: Dimensions of Selection Bias in Media Coverage of Washington Demonstrations, 1982, 1991." *American Sociological Review* 61:478–99.

McCarthy, John D , and Mayer N Zald. 1973. *The Trend of Social Movements in*

*America: Professionalization and Resource Mobilization.* Morristown, N.J.. General Learning Press.

———. 1977. "Resource Mobilization and Social Movements: A Partial Theory." *American Journal of Sociology* 82.1212–41.

Marwell, Gerald, and Pamela Oliver. 1984. "Collective Action Theory and Social Movements Research." *Research in Social Movements, Conflict and Change* 7 1–27.

Mehrlander, Ursula. 1994. "The Development of Post-War Migration and Refugee Policy." Pp. 1–14 in *Immigration as an Economic Asset: The German Experience,* edited by Sarah Spencer. Stoke-on-Trent: Trentham.

Meyer, David S., and Sidney Tarrow. 1998. *Social Movement Society.* Lanham, Md.: Rowman & Littlefield.

Miller, David L. 1985 "Migrations." Pp 159–70 in *Introduction to Collective Behavior,* edited by David L. Miller. Belmont, Calif: Wadsworth.

Mueller, Carol. 1997. "International Press Coverage of East German Protest Events, 1989." *American Sociological Review* 62.820–32.

Mushaben, Joyce Marie. 1984 "Swords to Plowshares: The Church, the State and the East German Peace Movement." *Studies in Comparative Communism* 27.123–35.

Naimark, Norman M. 1992. " 'Ich will hier raus': Emigration and the Collapse of the German Democratic Republic." Pp. 72–95 in *Eastern Europe in Revolution,* edited by Ivo Banac. Ithaca, N.Y.. Cornell University Press.

Oberschall, Tony. 1996. "Opportunities and Framing in the East European Revolts of 1989." Pp 93–121 in *Comparative Perspectives on Social Movements,* edited by Doug McAdam, John McCarthy, and Mayer Zald. New York: Cambridge University Press.

O'Donnell, Guillermo. 1986. "On the Fruitful Convergences of Hirschman's 'Exit, Voice, and Loyalty' and 'Shifting Involvements': Reflections from the Recent Argentine Experience.' " Pp. 249–68 in *Development, Democracy, and the Art of Trespassing: Essays in Honor of A.O. Hirschman,* edited by Alejandro Foxley, Michael S. McPherson, and Guillermo O'Donnell. Notre Dame, Ind.: University of Notre Dame Press.

Oliver, Pamela E., and Gerald Marwell. 1992. "Mobilizing Technologies for Collective Action." Pp. 251–73 in *Frontiers in Social Movement Theory,* edited by Aldon Morris and Carol M. Mueller. New Haven, Conn.· Yale University Press.

Olzak, Susan. 1989. "Analysis of Events in the Study of Collective Action." *Annual Review of Sociology* 15:119–41.

Opp, Karl-Dieter, Peter Voss, and Christiane Gern. 1995. *Origins of Spontaneous Revolution: East Germany, 1989.* Ann Arbor: University Michigan Press.

Pagnucco, Ron. 1995. "The Comparative Study of Social Movements and Democratization: Political Interaction and Political Process Approaches." *Research in Social Movements, Conflict and Change* 18:145–83

———. 1996. "Social Movement Dynamics during Democratic Transitions: A Synthesis of Political Process and Political Interactionist Theories." *Research on Democracy and Society* 3:3–38.

Philipsen, Dirk. 1993 *We Were the People: Voices from East Germany's Revolutionary Autumn of 1989.* Durham, N C.: Duke University Press.

Pierard, Richard V. 1990. "Religion and the East German Revolution." *Journal of Church and State* 32·501–9.

Piven, Frances Fox, and Richard A. Cloward 1992 "Normalizing Collective Protest." Pp. 301–25 in *Frontiers in Social Movement Theory,* edited by Aldon Morris and Carol M. Mueller. New Haven, Conn.: Yale University Press.

Pond, Elizabeth. 1993. *Beyond the Wall· Germany's Road to Unification.* Washington, D.C.: Brookings Institution

Ramet, Pedro. 1984 "Church and Peace in the GDR." *Problems of Communism* 33: 44–57.

Rokkan, Stein. 1975. "Dimensions of State Formation and Nation-Building: A Possible Paradigm for Research on Variations within Europe." Pp. 562–600 in *The Formation of National States in Western Europe,* edited by Charles Tilly. Princeton, N.J.: Princeton University Press.

Rucht, Dieter. 1990. "The Strategies and Action Repertoires of New Movements." Pp. 156–95 in *Challenging the Political Order: New Social and Political Movements in Western Democracies,* edited Russell Dalton and Manfred Kuechler. Cambridge: Polity Press.

Rucht, Dieter, Ruud Koopmans, and Friedhelm Neidhardt 1998. *Acts of Dissent: The Study of Protest in Contemporary Democracies.* Berlin: Sigma.

Schopflin, George. 1993. *Politics in Eastern Europe.* Cambridge, Mass.: Blackwell.

Sewell, William, Jr. 1990. "Collective Violence and Collective Loyalties in France: Why the French Revolution Made a Difference." *Politics and Society* 18:527–52.

Sharlet, Robert. 1978. "Dissent and Repression in the Soviet Union and Eastern Europe: Changing Patterns since Krushchev." *International Journal* 33:763–95.

Skilling, H. Gordon. 1966 "Interest Groups and Communist Politics." *World Politics* 18:435–51.

———. 1983. "Interest Groups and Communist Politics Revisted." *World Politics* 36: 1–27.

Smith, Jackie, and Ronald Pagnucco. 1992. "Political Process and the 1989 Chinese Student Movement." *Studies in Conflict and Terrorism* 15:169–84.

Snowden, Lynne L 1990. "Collective versus Mass Behavior: A Conceptual Framework for Temporary and Permanent Migration in Western Europe and the United States." *International Migration Review* 24:577–90.

Steinberg, Marc W. 1995. "The Roar of the Crowd: Repertoires of Discourse and Collective Action among the Spitalfields Silk Weavers in Nineteenth-Century London." Pp. 57–88 in *Repertoires and Cycles of Collective Action,* edited by Mark Traugott. Durham, N.C.: Duke University Press.

Tarrow, Sidney. 1989. *Democracy and Disorder.* New York: Oxford University Press.

———. 1993. "Modular Collective Action and the Rise of the Social Movement: Why the French Revolution Was Not Enough." *Politics and Society* 21:69–90

———. 1994 *Power in Movement: Social Movements and Contentious Politics,* 1st ed. New York: Cambridge University Press.

——— 1996. "The People's Two Rhythms: Charles Tilly and the Study of Contentious Politics." *Comparative Studies in Society and History* 38:586–600

———. 1998. *Power in Movement: Social Movements and Contentious Politics,* 2d ed New York· Cambridge University Press.

Tilly, Charles. 1978. *From Mobilization to Revolution.* Reading, Mass.: Addison-Wesley.

———. 1979. "Repertoires of Contention in America and Britain, 1750–1830." Pp 126–55 in Mayer N. Zald and John D. McCarthy, eds. *The Dynamics of Social Movements.* Cambridge, Mass.: Winthrop Publishers.

——— 1983. "Speaking Your Mind without Elections, Surveys or Social Movements." *Public Opinion Quarterly* 47:461–78.

———. 1984. "Social Movements and National Politics." Pp. 297–318 in *Statemaking and Social Movements,* edited by Charles Bright and Susan Harding. Ann Arbor: University of Michigan Press.

———. 1986. *The Contentious French.* Cambridge, Mass.: Harvard University Press.

———. 1995a. "Contentious Repertoires in Great Britain, 1758–1834." Pp. 15–42 in *Repertoires and Cycles of Collective Action,* edited by Mark Traugott. Durham, N.C.: Duke University Press.

———. 1995b. *Popular Contention in Great Britain, 1758–1834.* Cambridge, Mass.: Harvard University Press

Thomas, Daniel. 1997. "Norms and Change in World Politics: The Helsinki Accords,

Human Rights and the Demise of Communism." Doctoral dissertation. Cornell University.

Torpey, John C. 1992. "Exodus and Opposition in the East German Transformation, 1989–90." *German Politics and Society* 26:21–42.

———. 1995. *Intellectuals, Socialism and Dissent: The East German Opposition and Its Legacy.* Minneapolis· University of Minnesota Press.

Traugott, Mark. 1995. "Barricades as Repertoire: Continuities and Discontinuities in the History of French Contention." Pp. 43–56 in *Repertoires and Cycles of Collective Action,* edited by Mark Traugott. Durham, N.C.: Duke University Press.

United States Committee for Refugees. 1999. World Refugee Survey, 1998.

Wendt, Hartmut. 1991. "Die deutsche-deutschen Wanderunger—Bilanz einer 7 40 jahrigen Geschichte von Flucht and Ausreise." *Deutschland Archiv* 24:386–95.

White, Stephen, John Gardner, and George Schopflin. 1987. *Communist Political Systems,* 2d ed. London: Macmillan.

Woods, Roger. 1986. *Opposition in the GDR under Honecker, 1971–85.* London: Macmillan Press.

Zald, Mayer N., and John D. McCarthy. 1987. *Social Movements in Organizational Society.* New Brunswick, N.J.: Transaction Books.

# The Talk and Back Talk of Collective Action: A Dialogic Analysis of Repertoires of Discourse among Nineteenth-Century English Cotton Spinners[1]

Marc W. Steinberg
*Smith College*

This article offers a critique of framing perspectives on collective action discourse and an alternative *dialogic* approach. The argument set forth is that the latter sees collective action discourse as a joint product of actors' agency and discourse dynamics, including its multivocal nature. Such discourse is a joint product of challengers' rational actions and the constraints of the discursive field. Challengers seek to appropriate and subvert the dominant discourses that legitimate power, creating discursive repertoires. To illustrate this, the contentious actions of English cotton spinners in the 1820s and 1830s are analyzed. The spinners produced a discursive repertoire drawing on mill owners' dominant discourses.

Culture is once again squarely on the agenda of social movement analysis from a variety of analytic perspectives, examining an array of issues, including identity, media, ideology, and cognition (Jasper 1997; Johnston and Klandermans 1995; Taylor and Whittier 1995). While the range of analytic perspectives is large, the framing perspective dominates research informed by political process and resource mobilization models, as attested by recent writing (Johnston and Klandermans 1995; Klandermans 1997; McAdam, McCarthy, and Zald 1996; McAdam, Tarrow, and Tilly 1997). Frame analysis focuses on the social construction of meaning by social movement activists and organizations and on the media and the reactions between framers and potential supporters (Snow and Benford

1992; Gamson 1988*a*, 1992*a*, 1992*b*; Gamson and Meyer 1996; McCarthy, Smith, and Zald 1996).[2]

In this article, I engage what I see as several problematic features of the framing perspective and offer an alternative. I argue that rooted in an undertheorization of discourse processes is a series of linked problems concerning the theoretical depiction and reified empirical analysis of collective action frames and the framing process. As an alternative, I propose a *dialogic* analysis of collective action and social movement discourse. A dialogic analysis focuses attention on the production of meaning as essentially contested collective action that is motivated both by group conflict *and* the internal dynamics of discourse itself.

## FRAME THEORY

The most central work on framing theory has been elaborated by David Snow, Robert Benford, William Gamson, and their associates. Snow and Benford define a frame as "an interpretive schemata that signifies and condenses the 'world out there' by selectively punctuating and encoding objects, situations, events, experiences, and sequences of action in one's present or past environment" (1992, p. 137). Frame analysts describe framing as the process of creating the ideational elements of persuasive communication essential for both the mobilization of consensus prior to collective action and as the cognitive process necessary for orienting and sustaining collective action (Benford 1997, p. 410; Gerhards and Rucht 1992, pp. 578–84; Klandermans 1997, pp. 38–43; Snow and Benford 1992, p. 136). They generally depict movement activists or social movement organizations (SMOs) as strategic creators of frames that provide a compelling sense of injustice and the collective identities for the protagonists and their targets. Frames offer a diagnosis and prognosis of a problem and a call to action for its resolution (McAdam, McCarthy, and Zald 1996, p. 6; W. Gamson 1995, p. 90; Snow and Benford 1988, p. 199).[3]

Framing theorists maintain that the construction of frames is situationally sensitive, is keyed to interactive processes, and occurs in a recursive relationship with the dynamics of collective action and the broader cycles

---

[2] Frame analysts have devoted less systematic attention to the role of governmental authorities in these processes, though see Capek (1993).

[3] McAdam, McCarthy, and Zald suggest that framing has five principal facets for analysis: (*a*) the ways in which they are used as cultural tool kits by activists, (*b*) its strategic dimensions for collective action, (*c*) contests between challengers and authorities over frame meanings, (*d*) the ways in which the media are implicated in these contests, and (*e*) the impact that framing has in modifying the cultural tool kits available in the more general culture (1996, p. 17).

of protest that shape other social movement processes (Hunt, Benford, and Snow 1994, pp. 191–92; Klandermans 1988, p. 176; Klandermans 1992, pp. 82, 99). Consensual unity among a group concerning the contents of social movement frames is thus variable (Benford 1997, p. 422; Snow et al. 1986, pp. 477–78). Others have also observed that framing is bounded by the larger political culture or public discourse within which social contention develops (Donati 1992, p. 141; Gamson 1988a, pp. 221–22; Gamson 1992b, pp. 135–36; McAdam 1994, pp. 37–38). Generally, frame analysts argue that frames are derivative of ideologies, which they depicted as being more complex and encompassing systems of beliefs (Snow and Benford 1988, p. 205; Zald 1996, p. 262).

Snow and his colleagues suggest that successful framing depends upon the extent to which frames resonate with the potential understandings of adherents and sympathizers, which in turn is a function of their narrative fidelity, experiential commensurability, and empirical credibility (Snow and Benford 1988, p. 208; see also Gamson 1988b, pp. 167–68; Gamson 1992b, p. 135). Resonance can be enhanced by alignment processes that bridge the frame's messages to adherents' structurally congruent ideas, amplifying a particular value or belief, extending a frames ideological message to a wider pool of potential constituents and adherents, and sometimes transforming the contents of the frame to articulate a partly new message given current exigencies (Snow et al. 1986). Constructing frames that resonate with potential constituents and adherents thus partly results from using symbols and discourse familiar to the target population (Gamson 1988a, pp. 222, 242; Gamson 1988b, p. 166; Gamson and Modigliani 1989, p. 3). Snow and Benford (1992, p. 138) also have suggested that a master frame, a relatively stable configuration of ideational elements and symbols, operates as a kind of grammar for the articulation of more specific collective action framing processes within social movements.

KEY PROBLEMS

As a dominant mode of analysis, the framing perspective has spawned criticisms, many of them in-house.[4] It has been variously critiqued as lacking in conceptual precision in its delineation of constituent elements and processes, as reifying a signifying process, as being excessively voluntarist and nominalist, as having an analytic bias toward the elite production of frames, and as creating epistemological ambiguity between reality and its representations (Benford 1997, pp. 412–13, 418, 422; Ellingson 1995, p. 103; Gamson 1992a, p. 70; Jasper 1997, pp. 75–77; McAdam, McCarthy,

---

[4] I pursue this critique in more detail in Steinberg (1998).

and Zald 1996, p. 6; Zald 1996, p. 261; McAdam 1996, p. 340; Hart 1996, pp. 88, 95). I maintain that these and other problems arise from the perspective's lack of attention to what we can term the social semiotics of meaning production (Hodge and Kress 1988). This is manifest both in terms of a theory of meaning production in discourse generally and how meaning is ordered in narrative forms, though I will concentrate here on the former.[5]

As noted above, frame analysis depicts the framing process as a type of representational contest between actors. The focus of much of the writing from this perspective is on a somewhat reified *textual* level. Challengers and powerholders each proffer an ideological view of an issue through a discrete textual representation and seek to persuade others of its superior veracity. Thus, the discursive conflict focused on by current analyses is that which occurs *between* relatively modular and synchronic packages. Success depends on whether the arguments or expressed beliefs within the text have a logical coherency and congruity with the cultural understandings used by potential recruits and sympathizers to provide them with a real and compelling interpretation of the issue.

Implicitly, then, frames are depicted as relatively stable referential modes of representation. Frame analysts have been less interested in examining the production of meaning for the signifiers that compose a frame—the words, catch phrases, metaphors, and other symbols that are the vehicles of meaning. Moreover, they do not focus on the relations between these components both *within* a frame (or text) and between frames.[6] The underlying epistemology is that the transmission of meanings between actors is a largely uncomplicated process of sending and receiving messages. As Donati notes, frame analysis and many perspectives on political discourse "have adopted a positivist stance, considering language as a rational, denotational and neutral instrument; that is, as composed of words with specific and unique meanings" (1992, p. 157; see also Masson 1996, p. 78).[7]

---

[5] For a discussion of the role of narratives in social movements, see Polletta (1998). For discussions of the narrative construction of identity, see Somers (1994) and Somers and Gibson (1994) For theoretical approaches to the narrative analysis of hegemony and resistance, see Ewick and Silbey (1995).

[6] Even in their inaugural work on frame analysis, Snow et al. (1986) discuss the frame alignment process in terms of the manipulation of the "sentiments," "values," and "interests" contained in a frame and not the representational components themselves.

[7] Exemplary in this regard is the recent work of Gerhards and Rucht (1992) on the framing of a campaign against the International Monetary Fund (IMF). In analyzing what they term the "ideology of imperialism" frame used in these collective mobilizations, the authors represent the frame as a series of causal statements about the role of the IMF and World Bank in perpetuating global inequality, reducing them in fact to a schematic representation (1992, pp. 576–77). In doing so, however, they take as

In most current analyses, then, frame analysts ignore a foundational assumption of social semiotics, that is, that signifiers often can be interpreted in multiple, incongruent, and potentially divisive ways (Gottdiener 1995, pp. 19–22). In this sense, they do not wholly grapple with the *multivocality* of collective action discourse, the multiple meanings that can be conveyed and interpreted through any particular discourse. From a social semiotic perspective, meaning is produced in the *interaction* between social action and systems of signs (Hodge and Kress 1988, p. 6). As Carroll and Ratner cogently observe, "linguistic meaning may be said to have its sources both in semiotic relations among signs *and* in the contingent yet relatively durable relations that human practice establishes between signifiers and their referents" (1994, p. 15). On the one hand, this means that activists can never complacently assume that they can unproblematically convey a representation of an issue, since the words they use may be interpreted differently by their targets. As Roberto Franzosi notes, "There is never a single message uniquely encoded in a text; there are several messages ('a network of different messages') as decoded by different readers endowed with different 'intertextual frames' and 'intertextual encyclopedias,' and different reading codes" (1998, p. 533).[8] Frame analysts have paid insufficient attention to the ways in which such "networks of messages" themselves impose structured constraints on what can be represented. As I suggest below, the development of collective action discourses is both facilitated and limited by the ways in which claims and alternative visions can be represented within a larger discursive field. A social semiotic perspective focuses our attention on the inherent ambiguities in the representation of an issue and its resolution and the communication of such concerns between participants in contention. The multivocality of messages that actors (both powerholders and challengers) circulate leaves opportunity for claims and visions to be understood in ways that may be different from (and not necessarily congruent with) what they intend.

This lack of a basis in social semiotics leads to several underlying problems in frame analysis, one of which, particularly in the application of the framing perspective, comes in depicting the outcome of representational processes as reified units. As the metaphors of frame and package suggest, collective action discourse generally is conceptualized as a discrete set of bounded and linked issue statements. While frame analysts maintain that framing is a dynamic process of representation, in their analyses, they

---

unproblematic the meanings of many keywords that are used in the construction of the frame, such as "capitalism," "imperialist order," "third world," and even "IMF."

[8] Anne E. Kane raises this point from a more structuralist perspective (1997, pp. 250–51).

tend to depict frames as relatively stable meaning systems, akin to modular texts or maps, which can endure for long periods of time (and in the case of a master frame, over several lengthy social movements).[9] While framing analysts argue convincingly that collective action ideology is circumscribed, few if any such studies have demonstrated through the analysis of collective action discourse that its practitioners produce this discourse in such stable and structured forms.[10]

Depicting a frame as a discrete and clearly bounded map of meanings abstracts disparate and discontinuous discourse processes that ebb and flow over a cycle of collective action, highlighting a frozen moment in the course of action or outcome rather than the processes themselves. As Gamson and Meyer have noted, "The degree to which there are unified and consensual frames within a movement is variable and it is comparatively rare that we can speak sensibly of *the* movement framing. It is more useful to think of framing as an internal process of contention within social movements with different actors taking different positions" (1996, p. 283; see also Benford 1997, p. 422).[11] Activists, participants, and opponents are capable of reading different, divergent, and potentially contradictory meanings from frames. As Ellingson (1995, p. 107), Donati (1992, p. 161), and others have noted, meanings are susceptible to change over the course of a cycle of action.

If collective action discourse is not manifest in such stable and coherent modules, then we should be wary of characterizing it as representative of a discrete and distinctive oppositional subculture, as some analysts argue (Johnston 1991, pp. 49–50; Johnston and Klandermans 1995, p. 7; Mc-Adam 1994, pp. 45–46). While social movement discourse is certainly distinctive, Tarrow raises a poignant issue when he asks, "Why does it seem so difficult to construct truly oppositional symbols? " (1994, p. 125). The answer lies, I will argue below, in seeing culture itself as a terrain of struggle (Fantasia and Hirsch 1995, pp. 145, 158). Rather than positing exclu-

[9] Gerhards and Rucht's carefully diagrammed analysis of the collective action frame of the IMF protest (discussed above) is exemplary in this regard (1992, pp. 576–79). Benford, in noting the "descriptive bias" of framing analysis, lists over 50 discrete frames identified by the literature. As he observes, this is not the original theoretical meaning or methodological intent of the framing perspective, which argues that framing is an emergent process and that a multiplicity of frames can apply to a particular situation. However, he also notes that the metaphor itself is imprecise and is generally used in a monolithic manner (1997, pp. 413–16, 422).

[10] The exception in this regard is the area of media studies where Gamson and his associates, and others who have followed their lead, have demonstrated discrete structures in the media presentation of social and political issues (Gamson 1992a; Gamson and Modigliani 1989).

[11] In this sense, frame analysts have neglected the importance of what Goffman termed keying and rekeying (1974, pp 43–44, 88–89).

sive cultures and oppositional subcultures, we should analyze social movement culture more as an appropriation from a dominant culture (Voss 1993; Hart 1996).

A further problem is the degree of agency many frame analysts assume in their analyses of framing as strategic action. As McAdam, McCarthy, and Zald have recently emphasized, the conception of framing is of the "conscious strategic efforts by groups of people to fashion shared understandings of the world and of themselves that legitimate and motivate collective action" (1996, p. 6). Unquestionably, framing is strategic, but in the focus on calculation and persuasion, frame analysts have neglected the constraints and limits that discourse itself imposes on such agency. As Dominique Masson observes, discourse doubly constrains the way in which such agency is exercised: "Discourses limit both linguistic practices (the textual meanings that can be enunciated) and discursive practices as events (whether and how these events can occur)" (1996, p. 88).

A final issue, linked to the questions of agency, is that most current work on collective action discourse portrays frames as cultural resources and their production and dissemination as strategic action primarily orchestrated by movement activists and SMOs (Evans 1997, pp. 452–57; Johnston and Klandermans 1995, p. 8; McCarthy, Smith, and Zald 1996, p. 309). This conceptualization stems partly from the assumptions of rational action that underlie the resource mobilization and political process models. From this perspective, discourse is thus analyzed in terms homologous to material resources (Marullo, Pagnucco, and Smith 1996, pp. 2–3; McAdam 1994, p. 43; Williams 1995, p. 126).

However, this homology creates several conundrums. First, if collective action discourse is contextual, public, and emergent in the processes of mobilization and action, as most accounts suggest, then exercising control and distribution of it as a resource seems highly problematic. As even Fred Kniss, the most recent advocate of a modified version of this approach, notes, cultural resources are not commensurable, fungible, or divisible (1997, p. 135; Kniss 1996, p. 8).[12] Kniss also argues that such resources are more mobile, manipulable, volatile, and context dependent than material resources (1997, p. 136; Kniss 1996, pp. 9, 21; see also Jasper 1997, pp. 32, 46–47; Kane 1997, p. 254). Additionally, if discourse is deployed rationally, as are material resources, it is unclear how we can

[12] In their early work on injustice frames, William Gamson and his associates carefully differentiated resources—objects that have clearly specifiable units and are fungible—from what they termed know-how—skills necessary for effective mobilization and action (Gamson, Fireman, and Rytina 1982, pp. 23, 83, 86–87).

square this depiction with the social constructionist tenets of the framing perspective.[13]

These last problems of agency may be symptomatic of the fact that the underlying epistemologies of constructionist and rational actor accounts of collective action discourse are not easily reconciled. It is problematic to characterize social movement framing as both an exercise in the reality construction of genuinely held senses of injustice and identity, while simultaneously holding that activists and SMOs strategically manipulate and align frames to mobilize consensus. This can create an excessive voluntarism, vitiating the understanding of discourse as a stock of contested codes and meanings that impose boundaries on the ways in which people understand and represent their lives (Hart 1996, p. 88).

The discourse theory of the Bakhtin Circle and emerging literature in rhetorical social psychology offer a theory of discourse and ideology that resolves these conundrums. Additionally, they provide the basis for a reorientation of our analysis of cultural dynamics of social movements.

THE DIALOGIC ANALYSIS OF DISCOURSE

Frame analysis valuably focuses our attention on processes of ideological and cultural production for mobilization and action but undertheorizes the semiotic aspects of framing.[14] A dialogic perspective furthers our analyses of discourse dynamics, particularly in terms of its multivocality and the semiotic processes that underlie it.[15]

Dialogism focuses on discourse as an ongoing process of social communication.[16] It emphasizes the situational embeddedness of discourse and

---

[13] Polletta argues that this confusion stems from an underlying Parsonian division between instrumental and cultural action in the framing perspective (1997, p. 438).

[14] For works that highlight the ways discourse provides the structure by which actors fashion frames, see Ellingson (1995), Fine (1995), and Johnston (1991, 1995).

[15] As Sara Mills has recently observed, there is no simple definition for the term discourse (1997, p. 1). I take it to mean the process and product of socially situated and institutionally ordered ways people communicate their representations of lived and imagined realities. Discourse is most often, and most importantly, language in social use (though it can also be costumes, body language, pictures, and the host of other ways people communicate with one another). As Mills notes, "a discourse is not a disembodied collection of statements, but groupings of utterances or sentences, statements which are enacted within a social context, which are determined by that social context and which contribute to the way that social context continues in existence. Institutions and social context therefore play an important and determining role in the development, maintenance and circulation of discourses" (1997, p. 11; see also Macdonell 1986, pp. 1–3).

[16] This theory was developed by the Bakhtin Circle, named after Mikhail Bakhtin, a literature professor acknowledged as the intellectual center of this antiformalist group

its meaning as partly a product of social interaction and partly a product of how streams of language themselves interact. The analysis is thus both social and semiotic: social in that meaning is a function of the social interactions between people and the contexts in which these take place; and semiotic in that the languages themselves that people use and that are available to them to express their senses of the world limit of what can be expressed and understood (Bakhtin 1986, pp. 65–71, 122–27; Todorov 1984, pp. 41–54).[17] In the latter sense, dialogists emphasize that talk and texts people use in any specific exchange in part derive their meanings in relation to the wider sphere of talk and texts in ongoing communication. For Bakhtin, "The living utterance, having taken meaning and shape at a particular historical moment in a socially specific environment, cannot fail to brush up against thousands of living dialogic threads, woven by socio-ideologic consciousness around the object of the utterance; it cannot fail to become an active participant in social dialogue. And not all words for just anyone submit equally easily to appropriation . . . many words stubbornly resist, others remain alien" (1981, p. 276; see also Bakhtin 1986, p. 87; Volosinov 1986, pp. 21, 86).

What actors understand when they communicate with one another is thus a product of intersecting streams of communication produced in specific historical situations (Todorov 1984, pp. 60–61). With reference to challenging groups in collective action, for example, discourses of injustice noted by Gamson to be the fulcrum for framing take their meaning in relation to other discourses used in ongoing communication that depict power, difference, and hierarchy. In this sense, dialogism theorizes meaning production as purposeful but bounded by the larger field of relevant discourses in which meanings are produced, thus providing a coherent epistemology of both agency and its limiting structures within this cultural process (Shotter 1992, p. 15).

In addition, dialogism also emphasizes that discourse is essentially *multivocal* (Bakhtin 1981, pp. 291–92). Many words, phrases, and utterances do not have one unambiguous meaning but often have *multiple* meanings given their particular contextual use with other words, phrases, and utter-

---

(see Gardiner 1992; Morson and Emerson 1990; Clark and Holquist 1984). Many of their contributions were written in the 1920s and 1930s.

[17] For a similar appreciation of the contextual discourse strategies from a more pragmatist and interactionist approach, see the work of Gerald Platt and his associates on letters written to M. L King, Jr (Lilley and Platt 1994; Platt and Lilley 1994; Platt and Fraser 1998). Platt provides nuanced analyses of how both messages and social movement identification are constructed from within the particular sociocultural contexts of the writers, and how this is reflective of their multivocal capacity to rearticulate the meanings of movement discourse. For a more structuralist theorization, see Kane (1997, pp. 255–57).

ances and the knowledge and intentions of the actors involved. Meaning production is therefore a type of joint labor, a social production among and between actors that involves agreement, dissension, and ambiguity—sometimes minor, but at times considerable—which is always partly anticipatory, ongoing, and contains echoes of past usages (Burkitt 1998, p. 166; Shotter and Billig 1998, p. 16). And as Maria Shevstova reminds us, discourse is multivocal in part because it has socially diverse origins, the social divisions and hierarchies of the past and present being reflected and refracted in multiple interpretations (1992, p. 754; see also Barker 1997, p. 23).[18]

Often there is an ongoing struggle between actors trying to invest discourses with their preferred meanings, given their life experiences, situations, and their power to exert control over the meanings provided by words. As Bakhtin noted, "The word in language is half someone else's. . . . The word does not exist in neutral and impersonal language . . . but rather in people's mouths, in other people's contexts, serving other people's intentions: it is from there that one must take the word and make it one's own" (1981, pp. 293–94; see also Bakhtin 1984, pp. 121, 183; Collins 1999, pp. 139–40; Ponzio 1990, pp. 215–19, 253; Scott 1990, p. 177).

Bakhtin and many others since have emphasized that these struggles over meaning are often at the core of ideology in action. Discourse is ideological when the meanings it provides offer understandings about power, difference, and hierarchy that are claimed to be natural, accepted, or preferred. Dialogists have argued that discourse is ideologically saturated, particularly when it is used in group conflict. As Volosinov asserted, "The word is ideological phenomenon par excellence," and he tied discourse directly to contests of ideological power when he observed that, "The logic of consciousness is the logic of ideological communication, of semiotic interaction of a social group" (1986, p. 13; see also Bakhtin 1981, pp. 271–72; Medvedev and Bakhtin 1978, pp. 7–15). In this view, ideology does not exist outside of or prior to discourse but is created and structured in and through ongoing communication.

This emphasis has important parallels to the Gramscian theory of hegemony, which Snow and Benford reference in their early work as useful for understanding the framing process (Snow et al. 1986; Snow and Benford 1988; see also Carroll and Ratner 1996). Bakhtin argued that power-

---

[18] Goffman himself broaches this issue in a discussion of speech acts by noting transformations between situations and over time between the said, the normally understood, and what is understood on a particular occasion. However, his analysis is anchored in how situations and their definitions are responsible for such changes. He does not pursue this on the more semiotic level (Goffman 1981, pp. 64–65). Manning and Cullum-Swan in fact argue that Goffman's discussion of rekeying and reframing lacks a semiotic dimension necessary to understand such transformations (1992, p. 243).

holders can attempt hegemony through ongoing efforts to limit the way meaning can be structured within particular discursive terms, forms, and styles, as well as enforcing silence among the less powerful. They attempt to create in a "one-sided" exchange, more of a monologue than a dialogue. In practice, all communication has some attributes of monologue and dialogue to varying degrees, but much akin to Gramsci, dialogists emphasize that powerholders have the capacity to objectify meanings in discourse (Bell 1998, pp. 54–55). In such situations, as Masson notes, "Discourses also impose authorized 'ways of talking' about areas of knowledge or social practice on the broader institutional sites where they are hegemonic, and on the other actors wishing to intervene on these sites" (1996, p. 89). Such discourse becomes established as common sense and dampens the multivocal nature of social communication (Bakhtin 1984, pp. 189–90; Brandist 1996a, p. 103; 1996b, pp. 63, 70–71; Gardiner 1992, pp. 26–27, 164–65). This understanding of discursive hegemony adds to Snow and Benford's concepts of resonance, narrative fidelity, experiential commensurability, and empirical credibility. It suggests that which discourses are available to articulate injustice and its resolution, how they can be used in relations to other ways of talking about the world, and the degree to which powerholders and challengers can exert control over their meanings determine their experiential and empirical efficacy.

Dialogists argue that this process of objectivization and naturalization is partly produced and maintained through what Bakhtin termed "speech genres." Genres are "relatively stable types of utterance (with respect to content, linguistic style, and compositional structure) which in turn correspond to particular types of social activity. . . . Such genres mediate between sociopolitical and economic life on the one hand and language on the other" (Gardiner 1992, p. 81). They consist of the culturally and historically specific widely accepted sets of vocabularies, meanings, and rules of use, including social forms of interaction. Rather than conceiving a genre as a discrete system, dialogists identify it as "a kind of loose, multiform 'whole'" (Hoy 1992, p. 767). In this sense, as Ian Burkitt observes, the dialogic concept of speech genres is similar to Bourdieu's notion of generative schemes or structures of practice in that both refer to "a recursive pattern in social practices produced through socially instilled dispositions of agents that can be seen in abstract terms as a 'structure,' but which, in practice, subtly changes in each practical social context of activity" (1998, p. 165; see also Crossley 1999, p. 10).

Bakhtin and dialogists since have argued that in any period and society we can find a myriad of genres, from standard forms of casual street communication, to literary and scientific styles, and officialdom's highly formal language. All arise from and reflect back on institutionally organized

social processes. Powerholders, through institutional control and social standing, have the capacity to offer ideologically laden genres as popular common sense and as naturalized views of sociopolitical life (Bakhtin 1986, pp. 65, 79; Gardiner 1992, pp. 74, 81; Volosinov 1983, p. 116).[19]

Discourse is a conduit for hegemony, but discursive domination is prey to its own internal contradictions and thus is never complete. The multivocal nature of discourse provides an underlying instability in how it is interpreted, which can be highlighted by other discourses, traces of past usages that impinge on present meanings, or the cold realities of life that can demand an insufficient response through hegemonic discourse. Often, Bakhtin suggested, "a tense dialogic struggle takes place on the boundaries" of discourse (Bakhtin 1986, p. 143). Multivocality creates fragmentations and gaps in the production of a coherent and compelling common sense. It opens possibilities for contradictions of meaning and expression where once discourse enforced silence. Through these pressures on the solid facade of assumed meaning, challengers can see how other interpretations (and on rare occasions even other genres) might provide different and sometimes subversive ways of depicting their world, their struggles, and possible alternatives. As James Scott poignantly argues in the case of dominant discourses in what he terms the "public transcript," "We many consider the dominant discourse as a plastic idiom or dialect that is capable of carrying an enormous variety of meanings, including those that are subversive of their use as intended by the dominant. . . . For anything less than completely revolutionary ends the terrain of dominant discourse is the only plausible arena of struggle. . . . Any ruling group, in the course of justifying the principles of social inequality on which it bases its claims to power, makes itself vulnerable to a particular line of criticism" (1990, pp. 102–3).

It is this semiotic dynamic that informs a theory of how repertoires of contentious discourses are produced by powerholders and challengers in cycles of collective action. Rather than engaging in the wholesale process of pitting one discursive construction of social life and politics against a completely different alternative, challengers generally engage in a more piecemeal process of questioning certain meanings contained within a genre as the opportunity to problemize words and other representations presents itself.

---

[19] This has parallels with James C. Scott's concept of the public or official transcript, the dominant and public discourses through which powerholders exercise ideological control. As he observes, "The official transcript of power relations is a sphere in which power appears naturalized because that is where elites exert their influence to produce and because it ordinarily serves the immediate interests of subordinates to avoid discrediting these appearances" (1990, p. 87; see also p. 45).

## REFRAMING THE ANALYSIS OF CONTENTIOUS DISCOURSE

As outlined above, dialogism offers an alternative conception of discourse to that embedded in current work on framing. Rather than assuming communication as the sending and receiving of messages whose meanings are evident and unproblematic, it offers a model of discourse as a dynamic, conflict-ridden cultural terrain. Instead of initiating the analysis on a discrete series of frames (or a master frame), we should focus on how meanings within genres become open to reinterpretation and appropriation during conflict as collective actors use genres in combination to depict their understandings of justice, order, and equity. Many cultural and semiotic theorists argue that people cannot freely pick and choose such combinations and suggest there is a larger imposed structure that bounds this cultural action. They have conceptualized this structure in a variety of ways, but among the most widely used concepts is that of the discursive field (derived in part from the work of Pierre Bourdieu).[20] Such fields contain the genres that can be seen as contextually related when groups construct diagnoses, prognoses, and calls to action, and are partly structured in ongoing processes of hegemony (Crossley 1999, pp. 13–15; Gardiner 1992, pp. 74, 81). They are grounded products of ongoing social action (in this case contention). As Lyn Spillman argues, fields are a dynamic terrain in which meaning contests occur:

> Discursive fields do not simply determine meanings and values; rather, they form the limits within which cultural action occurs, and the tools for that cultural action. A discursive field forms the basis of all sorts of creative cultural work: it consists of the categories which make things mean, and not the meanings themselves. Within the discursive field, the particular meanings and values which emerge . . . exist in historically specific repertoires that we create and recreate: in this sense culture is contingent and creative. . . . But discursive fields are limits, as well as tools, because they are among the presuppositions which grant success or failure to mundane meaning-making. (Spillman 1995, pp. 140–41)

The construction of fields thus involves mutually recognized (though not always mutually accepted) cultural assumptions as to how and when a genre can be applied to a social situation, the extent to which it can relate to other genres, institutional rules for its use (especially in relation to other genres), and the relations between the actors themselves (particularly in terms of recognized hierarchies and power differences). As Nick

---

[20] Bakhtin discussed such intersections in terms of historically recurrent formations he termed chronotypes (Bakhtin 1981; Holquist 1990, pp. 108–15). However, Bakhtin's development of this concept for his "historical poetics" tended to focus on broad patterns of development within literary genres, and the concept remains relatively abstract and contradictory in his work.

Crossley observes, "each field presupposes certain abilities on behalf of those who engage in it and an appreciation of its rules, procedures and meaning structures" (1999, p. 14). The creation of fields thus has social, semiotic, and strategic dimensions that are all *relational* in nature: social in the sense that their use depends on some mutual appreciation of their applicability and interpretability and among actors, as well as a shared recognition of their capacity to use these genres; semiotic in that genres always achieve meaning in use with other genres; and strategic in that *within the limits* of the first two dimensions actors have reflective capacity and creative agency in meaning making (Bourdieu and Wacquant 1992, pp. 98–105; Burkitt 1998, p. 171).

Discursive fields are thus grounded in ongoing contention with institutional histories that define both the opportunities for and limits to the expression of injustice and its resolution through a type of fuzzy logic. As opposed to Bourdieu's use of the concept, however, I argue that the boundaries of a discursive field are never entirely fixed or clear. Because the structuring of genres has the three dimensions I noted above, how they are combined in a field, who is deemed as competent to use them, and when they are recognized as strategically appropriate is generally subject to some doubt and contention, often in modest but sometimes in central ways. Social movement analysts can provide us with many examples of these fuzzy boundaries from the "rights" discourses used in specific movement claims. For challengers seeking full inclusion in a polity, for example, rights claims often involve genres of law, citizenship, and nationhood within a widely recognized field in which these claims are raised. Yet challengers seeking full inclusion within a polity may also be structurally disadvantaged in many other ways, also involving contests over rights that do not readily fit into the generally recognized rules, practices, and meanings structures of a field that concerns political rights. The rights discourses involving access to employment and occupations, a "living wage," the practice of sexual preferences, or medical treatment can all raise issues of where the discursive field concerning rights claims begins and ends, as well as who is authorized and competent to define these boundaries.

The concept of a discursive field has some affinity with that of the master frame, since both provide for the identification of enduring discourses of contention within particular struggles, societies, and historical periods, but it also differs in a couple of critical respects. First, in their admittedly brief theoretical overview of master frames, Snow and Benford depict master frames as paradigmatic and generic systems of meaning: "master frames can be construed as functioning in a manner analogous to linguistic codes in that they provide a grammar that punctuates and syntactically connects happenings in the world . . . they provide the interpretive me-

dium through which collective actors associated with different movements within a cycle assign blame for the problem they are attempting to ameliorate" (1992, pp. 138–39). Their prime exemplar is the civil rights master frame initially developed by African-Americans in their struggles for equality in the 1950s and 1960s. In this definition, however, they elide the important distinction between signifiers—the words, phrases, and other vehicles of meaning—and the interpretations actors can plausibly make of them. Missing the multivocal nature of discourse, they pay insufficient attention to the ways that actors over time within a movement or actors from different movements can create distinctive meanings from the same words and phrases, meanings that might indeed stand in some tension to one another. Even though discourses can endure, dialogism pointedly reminds us that meanings can be altered significantly over time and between actors and situations. Second, the concept of the master frame is not specifically a *relational* understanding of meaning production as detailed by the concepts of genres and discursive fields, particularly in the social and semiotic senses discussed above. Contentious discourses are structured not so much by independent grammars of meaning, issue cultures, or some system of beliefs exterior to conflict; rather, they are determined by the ways challengers can combine genres, in particular, social struggles and the discursive fields often dominated by powerholders, in which this strategic action occurs.

The dialogic perspective also suggests an alternative metaphor for the depiction of collective action discourse. As opposed to a frame, we might more accurately conceive of movement- or action-specific discursive *repertoires,* akin to Tilly's notion of the collective action repertoire (Steinberg 1995*a,* 1998; Tilly 1995*a,* 1995*b*).[21] Tilly argues that "Repertoires are learned cultural creations, but they do not descend from abstract philosophy or take shape as a result of political propaganda; they emerge in struggle" (1995*b,* p. 42). In this account, repertoires are *relational* products of contention between challengers and powerholders, which limits both the strategic choice of performances as well as the conceptual mapping of possibilities for action. Moreover, "While contenders are constantly innovating . . . they generally innovate at the perimeter of the existing repertoire rather than breaking entirely with old ways. Most innovations fail and disappear; only a rare few fashion long-term changes in a form of

---

[21] Mooney and Hunt (1996) have recently suggested a similar modification of framing theory As they note, "A repertoire of interpretations suggests that movement participants (re)interpret and (re)construct systems of meaning already present in their life worlds" (1996, p. 179). However, they argue that social movement participants are drawing from "several persistent master frames to (re)construct their ideological claims" (1996, p. 179), whereas I find the concept of the master frame to be more problematic.

contention. Only very rarely does one whole repertoire give way to another" (1995b, p. 44).

Akin to Tilly's notion, we can see that a central part of the development of discursive repertoires is done interactionally with opponents and targets through a process of conflict. Within the discursive field defined both by the conflict and its institutional histories, actors draw on those genres through which they saliently and compellingly can depict a shared understanding of injustice, identity, righteousness for action, and a vision of a preferred future. How challengers combine genres to create a repertoire is thus an outcome of past conflict, their strategic deliberations, and how genres within a field limit their possibilities.

This interactive emphasis on repertoire creation suggests a shift in our understanding of how actors produce oppositional culture. As opposed to pitting two distinct frames—or a subculture and a dominant culture—against one another, dialogic analysis focuses more attention on the ways in which challengers seek to delegitimate hegemonic genres within a field while appropriating pieces to inflect it with their own subversive meanings (Steinberg 1998).[22] As I noted above, the multivocal nature of discourse provides the means for challengers to find gaps, contradictions, and silences in this taken-for-grantedness of hegemonic genres. By exposing these, challengers can inject alternative meanings to articulate their sense of injustice and moral authority for collective action (Barker 1997, p. 22; Carroll and Ratner 1994, p. 6). This process is piecemeal—a kind of war of position keyed to struggle—and often initially provisional, since challengers do not have the institutional bases or social standing to legitimize their oppositional meanings (Hunt 1990, p. 314).[23] Innova-

[22] In this conception, I also diverge from James Scott's concept of the "hidden transcript," which according to his theory of infrapolitics is where we find the counterideology of subordinates. Scott certainly acknowledges that "the frontier between the public and the hidden transcripts is a zone of constant struggle between dominant and subordinate—not a solid wall" (1990, p. 14). Yet he also maintains that "ideological resistance can grow best when it is shielded from direct surveillance" (p. xii), and that ideological resistance is best nurtured in sequestered sites. As opposed to Scott, it seems to me that a dialogic theory places more emphasis on the public sites and discourses of ideological struggles. The critical moments of delegitimation and appropriation are in the end public, for that is when the struggle is fully engaged. Perhaps it would be fair to argue that initial processes of appropriation go on behind the backs of powerholders through public communication but in perhaps furtive and shrouded ways that have some parallels to Scott's concept of the hidden transcript. As Scott himself suggests, "what permits subordinate groups to undercut the authorized cultural norms is the fact that cultural expression by virtue of its polyvalent symbolism and metaphor lends itself to disguise" (p. 158)

[23] As Masson emphasizes, "Discourses are not omnipotent nor agentless Rather, they draw their authority and social efficacy from the repeated utterance or performance of their elements by people speaking from particular socio-enunciative positions or institutions inscribed within a field of power relations" (1996, p. 76).

tion, as Tilly suggests in the case of collective action repertoires, often comes at the margins. Moreover, such innovation develops both because of the contours and contents of the discursive field in which the struggles occur and the strategic aims of challengers and their targeted audiences.

As challengers develop discursive repertoires through this process, they also peer into the actions and histories of other challenging groups and their struggles to see what is available within their discursive repertoires. As Snow and Benford argue in the case of the master frame, borrowing oppositional schema from other challengers is a tried and true method of articulating a group's own injustice claims. They focus on how challengers borrow master frames to provide a movement or collective action frame for claims (such as the example of the civil rights master frame) or add arguments and language in the hopes of aligning other groups with their cause (Snow et al. 1986, 1992). However, such borrowing also can be a way of developing a discourse where a field fails to provide the words, catchphrases, language, and other symbols needed to articulate a sense of just opposition and all that it entails. Providing a cogent example in his work on the rise of the British mental health movement in the early 1970s, Nick Crossley notes that in their initial protest actions mental health patients referred to their organization as a "union" and their protest action as a "strike" because at the time this class struggle language was the most accessible to them to characterize their challenge. Further, while they first identified themselves as "mental patients" in these early actions, by the mid-1980s with the rise of the feminist antiviolence movement, they began to characterize themselves as "survivors" (Crossley 1999, pp. 6–7). In neither case were they seeking to necessarily cast their struggles in terms of these other movements, nor were they seeking alignment with other groups. Rather, facing a dearth of ways of expressing opposition within the discursive fields of psychiatry and law, they appropriated familiar terms from other movements to develop their own repertoires. As he astutely argues, challengers have a self-reflexive understanding through which they find that "their movement and struggle is, for them, akin to other specific struggles which they are aware of and belongs to a more generic and general typification of 'movements' and 'struggles.' . . . This establishes a possibility for the trafficking and transference of repertoires" (1999, p. 7). Dialogically, these typifications allow challengers to borrow discourse from other fields to be able to articulate identities, grievances, and goals where there are gaps and silence in the discursive field in which they are fighting their own struggle.

Throughout this process of repertoire development, we can expect that discursive and collective-action repertoires are roughly keyed to one another and can be mutually stabilizing, though if they diverge, the cogency of one or the other can be held up for scrutiny. Moreover, from the multidi-

mensional perspective described above, we should expect that there is some congruity and reinforcement of networks of actors and discursive repertoires within a cycle of contention. Strong ties between groups or activists create the social relations that can foster the shared production of meanings within a particular repertoire and field, while weak ties are more likely to allow for divergent production of meanings.[24]

Developing discursive repertoires obviously has a strategic dimension. However, in contrast to frame theory, a dialogic analysis suggests a way of squaring both its constructionist and instrumental aspects. First, while challengers consciously seek to appropriate and transform hegemonic genres, they are always partly captive to the truths these genres construct. Their version of the truth, after all, is in part predicated on the veracity of that which they appropriate, and social cognition is structured within discourse. Challengers and powerholders thus never stand completely outside the meanings imposed by dominant genres and fields, and framing is never simply a strategic process. Second, as an intersubjective process of meaning production for both the challenging group and for potential sympathizers, discursive repertoires must establish a requisite moral integrity. Challengers cannot simply readily and instrumentally manipulate discourses for their own cynical ends, for this undermines the foundations of mutual understandings that explain the justice of their claims and actions to themselves and others.[25]

Finally, this alternative perspective offers an argument against viewing collective action discourse as a resource. If challengers generally remain partly captive within hegemonic genres, then it is problematic to characterize that which is partly bounding and constraining as a manipulable resource.[26] As I have already argued, viewing discourse as a resource and as a mediator of social reality rests on contradictory epistemological as-

[24] In their neo-Gramscian analysis of seven social movements in Greater Vancouver, Carroll and Ratner (1996) establish that such strong ties facilitated the shared meanings for a widely used "political economy" master frame.

[25] I discuss additional aspects of movement scope and temporality in Steinberg (1998). Rather than a master frame losing its potency for a movement, as Snow and Benford (1992) theorize, we can alternatively conceive of a growing dispersal of meaning over time. As group networks and numbers multiply, situated meanings for signifiers, vocabularies, and so on in a repertoire might as well. Cracks in assumed commonalties of meaning can then become increasingly exposed. This might be particularly true in terms of discursive constructions of "we" in identity movements, as noted by Joshua Gamson (1995, p 397).

[26] As Diane Herndl observes in the case of feminist dialogics, the dilemma of critique is that challengers are never situated wholly outside of the dominant discourse. A feminist critique in the dialogic sense means a positioning that is partly "not-masculine" rather than something that can be unequivocally characterized as feminism (1991, pp. 16–17).

sum ptions. Additionally, we can now see that discourse can never be exclusively controlled nor even presumed to have a stable value or utility, as do material resources. Indeed, this is precisely why hegemonic discourse can never be secure, nor why discursive repertoires are in and of themselves likely to be the fulcrums of success or failure. As Peter Hitchcock poignantly observes, "Strategies of language use, conscious or otherwise, are not in themselves the means to transform society," though they can play an important mediating role in facilitating social change (1994, p. 8).

Having outlined both a dialogic theory of discourse and its application to cultural processes in collective action, I now briefly illustrate the perspective through a look at the discursive repertoires of the cotton spinners of southeastern Lancashire in the later 1820s and early 1830s. I collected texts for analysis from all of the Manchester and Stockport commercial newspapers that covered the event (these being the principal cities in the strike region with presses), all newspapers published by the cotton spinners and their allied union associations for the period, tracts written by popular political economists specifically targeted at factory workers, and pamphlets, broadsides, placards, and summaries of public speeches concerning these conflicts available from archival sources.[27] All texts were in some way circulated in the public sphere, that is, none were for private union, employer, or state communication. The vast majority of the texts were directed toward striking spinners themselves or their working-class compatriots (many of whom were other factory workers idled by the strike). A few texts, such as the series of 16 weekly pamphlets produced by the Manchester spinners during their prolonged strike in 1828, also spoke to the general public to win support and responded to factory owners' assertions as well. However, as I have argued above, dialogic analysis suggests caution in focusing solely on the strategic targets of addresses: actors' intentions do not preclude them from communicating with others whom they do not target, and in ways they do not envision.

From this collection, I extracted those texts that provided some justification—moral, social, or economic—for the stance taken by each of the principals. Many texts from commercial newspapers provided brief accounts focusing on material demands of each side, issues of pounds and pence, or the specific actions taken in mobilization for or during strike

[27] As Roberto Fransozi (in press) astutely argues, all historical research based on this type of textual analysis suffers from sampling and validity biases. I can make no claims that this collection of texts is properly representative of all publicly circulated texts for these conflicts. However, all sources report on public events or are public commentary. I refer all readers interested in how dialogic analysis is used for complete records of discursive conflict to Barker (1997) and Collins (1999). Both papers analyze dialogic processes of conflicts based on transcriptions of contentious meetings.

actions, without provided spinners' or factory owners' justifications of claims or positions. In the case of the factory owners in particular, I have had to rely more heavily on more general public speeches and tracts produced by allied newspaper editorialists and political economists who directed their work at the spinners during this period. While the factory owners had confederations to oppose spinners' demands, particularly in Ashton where, as I detail below, 52 firms signed a pact binding themselves to a stated wage maximum, this group issued few public pronouncements and had no public spokespeople (with the possible exception of the factory owner and civic luminary Charles Hindley, whom I discuss below). However, the local commercial newspapers, widely read and accessible by all groups in the region, operated as the informal conduits of the factory owners. Whenever possible, my use of tract literature comes from reprinted selections from these papers, which frequently provided such excerpts as a didactic feature or quoted such tracts to bolster their own editorials on the general state of the cotton industry and the strikes more specifically. The spinners had both their own union papers and more visible union leaders, and both more directly and clearly annunciated the underlying logic of injustice that motivated their actions. To analyze their discursive repertoire, I have relied largely on these newspapers and supplemented them with pamphlet literature and accounts of meetings produced during strikes.

My mapping of the discursive field is thus a grounded product of the dialogues between the spinners and employers and their sympathizers that emerged over the course of contention. It focuses on recurrent discourses keyed to (though not limited by) the hegemonic genres in the discursive field. Within limited space, the analysis focuses on the structure of the discursive field, the hegemonic genres that importantly formed its boundaries, the principal discourses in repertoire developed by the spinners during contention, and the ways in which this repertoire was a relational product of discursive contention.

## THE FACTORY OWNERS' AND SPINNERS' INTERPRETIVE REPERTOIRES IN A CYCLE OF PROTEST

From the middle of the 1820s through the beginning of the following decades, the cotton spinners of southeast Lancashire found themselves locked in a cycle of strike activity that pitted them against an increasingly powerful group of factory owners (Cotton 1977; Hall 1991; Kirby and Musson 1975; Sykes 1982). The spinners were the elite skilled male workers in cotton spinning factories whose workforces were disproportionately comprised of unskilled women and children (Catling 1970; Freifeld 1986; Huberman 1986; Lazonick 1979, 1981; Valverde 1988). Thus, at issue in

these contests were not only the piece rates but their collective identity as the male vanguard of the factory and their status in the communities in which they lived.

The spinners were the most highly organized of all groups of factory workers and among the most unified trades in all of Britain. By 1828, they had founded a union that formally united all local "combinations" through the United Kingdom, establishing a representative structure and a series of union papers (Catling 1970, pp. 148–49; Doherty 1829; Fowler and Wyke 1987, pp. 14–35; Turner 1962). In 1830, their secretary, John Doherty, led a campaign to establish a union of all trades throughout the country that drew considerable support from diverse trade groups in the industrial North and the Midlands (Kirby and Musson 1975).

Starting in 1825, with major strikes in Chorley and Stalybridge (two cotton towns in the hinterlands of Manchester), spinners sought to protect the deterioration of their piece rates and their control of the labor process. These strikes included highly contentious, protracted, and sometimes violent actions in Stockport in 1828 and Manchester in 1829 and culminated the cycle in a massive strike of 52 firms in the towns of Ashton-under-Lyne and Stalybridge in the winter of 1830–31 that idled some 20,000 workers (Cotton 1977, pp. 211–34; Hall 1991, pp. 101–12; Kirby and Musson 1975, pp. 31, 43, 57–58, 119–38; Tufnell 1834, pp. 18–19; Steinberg 1999, chaps. 8, 12). These strikes were generally prompted by factory owners' announced piece rate reductions, and in the case of Manchester, the underlying threat of the use of female and young male labor (British Public Record Office, Home Office Papers [hereafter HO], ser. 40, box 27, ff. 163–64, G. R. Chappell to Peel, Oct. 23, 1830; f. 342, Foster to Peel, Nov. 13, 1830; *Manchester Guardian* [hereafter *MG*], Dec. 18, 1830; *Manchester Times* [hereafter *MT*], Dec. 25, 1830, Feb. 26, 1831; *United Trades' Co-operative Journal* [hereafter *UTCJ*], Apr. 3, 1830; Doherty 1829, p. 51; Freifeld 1986, p. 334). A crisis of accumulation in the cotton industry in the latter 1820s put a severe squeeze on capitalists' profits, in what had earlier been an industry of rapid growth and fast fortunes (Baines [1835] 1966, p. 359; Howe 1984, p. 25; Sykes 1982, pp. 23, 123; British Parliamentary Papers [BPP] 1833, p. 652). Factory owners reacted to trade depressions and market gluts by organizing against increases in the price of raw cotton, opting for vertical integration of spinning and weaving, making pacts to work their mills at reduced hours, intensifying the labor process, and imposing piece-rate reductions on their workforces (Francis Place Collection of Newspaper Clippings and Pamphlets [hereafter PC], set 16, vol. 2, "Cotton," f. 73; *MG*, Dec. 11, 1830; *Stockport Advertiser* [hereafter *SA*], Jan. 29, Feb. 5, 1830; *Wheeler's Manchester Chronicle* [hereafter *WMC*], Feb. 6, 1830; *MT*, Sept. 12, 1829; BPP 1831–32, p. 430; BPP 1833, p. 685; BPP 1837–38, p. 252, 271–72; Baines 1835, p. 381; Catling 1970,

p. 54; Huberman 1986, p. 993). These reductions galvanized entire factory towns behind the spinners' efforts, for a severe depression in the late 1820s had left thousands of workers in these districts unemployed and thread-bare. At stake were not just wages but cohesive communities and a quo-tidian working-class culture that had grown up in the shadow of the facto-ries.

The spinners thus found themselves pitted against a fairly unified and ascendant group of capitalists who distinguished themselves by a distinc-tive "middle-class" culture.[28] Across a broad spectrum of ideas, social ac-tivities, and family histories, the region's mill owners fashioned a new bourgeois order. As Anthony Howe, a historian of the group, has observed, "The distinctiveness of the cotton masters, as a group, was the product not only of their particular social and economic formation, but also of their relationship to the dominant aristocracy and the nascent working class. . . . As a middle class—in terms of wealth, power, status, and cul-ture—the cotton masters developed a separate identity and their own spe-cific organizations" (1984, p. v).

A central facet of this social organization and collective identity was a hegemonic vision of the new order they saw themselves erecting. Drawing from genres of popular political economy, liberal politics, and notions of piety anchored in dissenting faiths, these master manufacturers sought to impress this vision not only on their middle-class peers and their aristo-cratic associates but on the mass of working people in the region as well.[29] As used in moral, political, and economic analyses and arguments by the manufacturers and their allies, such genres overlapped within the discur-sive field in which they attempted hegemony, not only in the manufactur-ing sector, but in political and social life as a whole. Importantly, this field was not a seamless, prescriptive set of strictures of what could be enunciated or argued within the ranks of elites and pundits. There were certainly contrarians who used theological discourses who argued against the principles articulated by political economists, for example (Horne 1990). However, the field exerted a strong influence on how genres could be conceived as linked in making claims and moral arguments both among powerholders and challengers.

By the 1820s, a loose set of discourses was being fashioned into a hege-monic repertoire. This became algorithmic for social, economic, and politi-cal analysis in the regions' commercial (middle-class) newspapers, includ-ing the *Manchester Guardian*, *Manchester Mercury*, *Manchester Times*,

---

[28] This culture was not simply uniform. As John Seed (1982, 1986), R. J. Morris (1990), Theodore Koditschek (1990), and others have noted, there were divisive squabbles among merchant and industrial groups, particularly concerning religion.

[29] I pursue this line of analysis in detail in Steinberg (1999), chaps. 9 and 10.

and *Stockport Advertiser* and the more highbrow publications of the Literary and Philosophical Society of Manchester, which had an important coterie of manufacturers and financiers.[30] Particularly by this period, capitalists within this industrial heartland had succeeded to a remarkable degree in setting the boundaries of the discursive field within which contention occurred. Reflecting on the course of class conflict, the Manchester Trades' Committee (a union group) observed in 1832 that "The exhortations of employers, have, for the last twenty years, kept up a perpetual warfare between themselves and the operatives" (*Union Pilot and Cooperative Intelligencer* [hereafter *UPCI*], March 10, 1832). Having been immersed in these battles, committee members were acutely aware of the master manufacturers' well-developed discursive repertoire in this "perpetual warfare." One of the dominant facets of the capitalists' repertoire for the spinners who battled them was construction of the economy and labor relations through the genre of popular political economy.

Popular political economists portrayed the economy as a great fiscal system subject to its own natural laws, with each class having an appropriate place (Berg 1980; Claeys 1985, 1987; Goldstrom 1985). As self-proclaimed disciples of Smith, they laid the foundations for this theoretical edifice in a labor theory of value. One of their more prolific pamphleteers, Charles Knight, in a tract devoted to the follies of unions written on the heels of the great strike in Ashton and Stalybridge, readily acknowledged that "All property is the result of labor" (1831, p. 7). Yet in the same breath, he quickly noted that "the accumulation of property is that which makes a rich [*sic*] community superior in all accommodations of life to the poor" (1831, p. 7). Other such writers effusively noted the virtues of capital in this process of creating a society of benevolence and comfort. According to Thomas Hopkins, a popularizer who gave a series of public lectures in Manchester during the series of spinners' strikes, the fuel of the system was capital, "that portion of wealth which is used for the purpose of producing new wealth" (*MG*, Oct. 3, 1829). Capitalists and workers existed in a mutually symbiotic relationship through the use of capital in the great circuit of consumption and production it animated:

> A profit was paid to the owner of capital, on account of the productive power it possessed; and . . . the use of the capital increased the produce of the labourer to a greater extent than the amount of profit taken by the capi-

---

[30] As Fetter (1965) observed, three of the four prominent national reviews (the contemporary equivalent to national magazines)—the *Edinburgh Review, Quarterly Review,* and *Westminster Review*—were also clearly espousers of some version of political economy, though the Tory *Quarterly Review* parted company with the other three on particular issues. These reviews were critical in establishing the discursive field for political debate.

> talist for the labourer, and . . . the latter was consequently always benefited by the use of capital, although he had to pay the profit of it. . . . In a state of society where the whole of the capital was owned by a separate class, that class had to set the labourers to work with it, and . . . the rates of wages and profits were determined by the laws of nature, independent of the wills or wishes of either party. (*MG*, Oct. 3, 1829)

One of these laws, of course, was that the capitalist must obtain a certain return on his property. For both mill owners and political economists, the necessity and virtue of profits above all else was self-evidently given. Such was the position of Ashton mill owner Charles Hindley, a stalwart free trader, self-described champion of the laboring classes, and future member of parliament (M.P.) for the town. At a meeting of mill owners in Manchester in January 1830, called to combat the rising price of raw cotton, he employed this logic to flatly state their limited recourse: "The difference in price must come out of the pockets of the workmen. He lamented that result, but it was inevitable. . . . All capital was invested for the sake of profit, and if it did not yield that, the spinners must either stop their mills or lower their rate of wages" (*MT*, Jan. 30, 1830). In a tract written in the wake of the Ashton spinners' strike, Hopkins reiterated the importance of profits to this system in a section that was part reminder and part rebuke: "Let the journeymen weavers and spinners in any district consider what would be their situation if the capital which is now employed in setting them to work, were to be entirely withdrawn, and then they will have some conception of the important service that profit renders them in preventing such withdrawal" (1831, p. 8).

On the flip side of the equation, a cardinal assumption of manufacturers and political economists was that labor was no more nor less a commodity whose price was governed by these natural laws. As correspondent "H. H." observed in his letter to the *Guardian*, commenting on a recent address by the general secretary of the spinners' union, "labour is just as tangible and marketable as any other commodity, and will, in the same way, find its real value and level" (*MG*, Dec. 19, 1929). As we shall see in the analysis of the spinners' discursive repertoire, the multivocal nature of this and other terms in the discourse created substantial potential for appropriation.

Derivative from the above law, of course, was the axiom that working-class mobilization and action to raise wages were in the long run futile and misplaced efforts, because they could not change the total amount available for the wage fund. As the *Guardian* commented during the Ashton strike, "Nothing is more certain than that turns-out never did and never can produce any permanent advance in the rate of wages. This must necessarily depend on the amount of capital applicable to the employment of labour, compared with the number of labourers wanting work" (*MG*,

Dec. 18, 1830). Whether emphasizing Malthusian dictates or focusing on the shifting dynamics of the market, the bottom line was that workers could only demand what the system would bear. As an enervated Stalybridge master manufacturer wrote to the *Stockport Advertiser* during the Manchester spinners' strike of 1829, "A temporary turn-out can do no good, there must either be a permanent reduction in the quantity manufactured, or prices cannot rise. Is it not, therefore, madness for the operatives to attempt to get what is out of every body's power to give them? Do they suppose their masters are made of money, and can give them as much as they demand? Trade is uncontrollable either by masters or men" (*SA*, Jan. 23, 1829). Within the discourse of political economy, agency over market forces was eviscerated, denying the manufacturers of culpability and the workers a rationale for collective action.

Of equal certainty was that whatever follies workers might pursue, the inviolable nature of the capitalist's property was to be maintained by force of law at all costs, not only for their benefit but for the society as a whole. It was at this point in the manufacturers' discursive repertoire that the discourse of political economy was wed to that of political liberalism. Under the rallying cry of free trade, factory owners and political economists constructed a vision of both polity and civil society centered on the sanctity of the individual (rather than the collective), negative rights, and a minimalist state. The author of *A Short Address to Workmen on Combinations to Raise Wages* (published not long after the end of the Ashton strike) broadly and unequivocally defined freedom, oppression, and justice in these terms: "The principles and laws of a government founded on justice ... must ensure to the individuals who compose a nation an entire security for personal freedom, and an unfettered use of their property, labour, and industry. With all unnecessary restriction on these two essential points, tyranny commences; every interference with these objects not sanctioned by law, or when sanctioned by law not founded on the public and general good, is oppression. These principles hold universally" (*MG*, Jan. 21, 1832).

The imagined community of the nation within this liberal discourse was one of a compact of individuals in which the collective good was defined as the mutual recognition of private interests. In a companion volume entitled *The Rights of Industry*, the author reminded working people that the defense of private property was thus by natural law synonymous with the very security of society itself: "The rights which are most open to attack from ignorant and designing persons, are the rights of property. Upon the upholding of these rights depend your own security, your own freedom, your own certainty of going steadily forward to the improvement of your condition. ... Nothing can destroy our ultimate peace and prosper-

ity but a violation of the great principles of natural justice, by which property is upheld for the benefit of all (*MG,* Dec. 3, 1831).[31]

Within this nascent hegemonic genre was an unremitting vision of progress carrying forth the ideal 18th-century notion of *doux commerce* as a civilizing force in all facets of society (Hirschman 1986). This is perhaps captured no better than by Charles Hindley, the Ashton factory owner whose storied public career as both M.P. and civic advocate led him to campaign for a host of liberal causes, including (perhaps perversely) factory reform legislation (Follows 1951). Speaking at a public dinner shortly before the Ashton strike in 1830 (in which his own workers were to participate), Hindley presented a broad image of unfettered trade as *doux commerce,* a system inextricably connected to and promoting of all liberal and religious causes championed by his fellow manufacturers.

> And in the name of common sense, which refuses to return to the destitution of a barbarous and savage age—in the name of reason, which proves the restrictive system to be absurd—in the name of experience, which has found it to be impossible—in the name of morality, which deprecates falsehood, and abominates slavery and war—and, above all, in the name of religion, which teaches us that God has made of one blood all the families of the earth—in the name, I say, of common sense, reason, experience, morality, and religion, I propose unrestricted free trade . . . "Free Trade all over the World." (Hindley 1841, pp. 15–16, 24)

Hindley's speech suggests the broad ideological terrain within which struggles between capitalists and workers were fought, one in which a dominant field mapped the contours and interconnections between social, economic, and political life. As we will see below, these hegemonic genres prompted a corresponding repertoire by the spinners throughout the cycle of collective action.

Finally, all of these genres mapped an understanding of gender relations onto social and economic life and, therefore, the conflicts as well. Within popular political economy, the implicit gendering of all actors was masculine.[32] This was both because of the constructions of gender and domesticity that were hardening into a discourse of public and private spheres in this period for all classes, and because all concepts of ownership and con-

---

[31] The *Manchester Guardian* similarly pronounced the imperative of protecting private property at the commencement of the Ashton strike· "But to the accumulation of capital, a peaceable state of society, and security to property, are absolutely and at all times essential. No country can prosper, no numerous body of labourers can long exist, without a full protection is given to that property which constitutes the fund for their maintenance" (*MG,* Dec 18, 1830)

[32] Carole Pateman (1988) has traced the lineage of this embedded masculinity in her analysis of the development of the Enlightenment thought on politics and economy.

trol within legal discourse were masculine (Hall 1992; Rose 1992, 1993). Women of course were a majority of the labor force in the cotton factories, and they were active supporters of the spinners during their strikes (many of course being related in one way or another to them). Yet the masculine coding of voice, power, and agency within all of these hegemonic discourses created a space of silence for women, which was reciprocally reinforced within the social relations of the factory, home, and community (a point to which I will return).

## THE SPINNERS' DISCURSIVE REPERTOIRE

As I noted in introducing this case, the spinners throughout the late 1820s and early 1830s engaged in a series of reactive actions to stave off piece-rate reductions. Much of their public claims making centered around the particulars of piece-rate variations between both regions and factories. Underlying these claims, however, were larger depictions of their labor and its value, their contributions to civil society and their collective identity within it, and the role of government in validating and protecting all of these. Given the contiguous genres available within the discursive field as partly mapped out by the manufacturers' hegemony, the spinners' repertoire varied to a certain extent with each particular campaign. Recurrent themes can be found undulating through these specific actions. From a dialogic perspective, we can see these themes emerging through a counter-hegemonic process, as the spinners developed their repertoire in interaction with and response to the hegemonic field. The field in which they developed their repertoire was significantly defined by the genres of powerholders, though it was also partly circumscribed by a vibrant campaign for Parliamentary reform that was concurrent with these strikes, which fostered a powerful political radicalism. Having overlapping sets of collective actors and clear issues of power and its discontents, and the social and semiotic relations between these campaigns, provided the potential for appropriating discourse. Thus, the spinners sought to both legitimize their claims within the discursive field of their conflicts and expose the power-laden nature of its representations through their depictions of economy, polity, and society. Their repertoire was shaped both by their own agency in interaction with their nemeses and targets and the contours of the discursive field that provided possibilities for meaning.

A principal part of this repertoire was developed through the appropriation of popular political economy. As Noel Thompson argues (1984, 1988), a nascent socialist economic discourse that emerged during the 1820s and 1830s was partly developed through a reshaping of political economy (hence the often-used label of "Smithian" socialism). However, other genres of economic discourse composed the genre assumingly available

to working-class actors such as the spinners. Indeed, a well-developed discourse of Owenite cooperative socialism was widely diffused by the late 1820s (Claeys 1987, 1989; Thompson 1984). There were several working-class periodicals available to the spinners, which emphasized Owenism, and indeed periodicals connected to the spinners such as the *United Trades' Cooperative Journal* and the *Voice of the People* contained favorable articles that discussed the principles of cooperation. However, cooperationist discourse, or indeed any other alternative economic discourse not appropriated and derived from popular political economy, was absent from the spinners' discursive repertoire, for perhaps three overlapping reasons. First, the spinners were engaged in a highly reactive campaign against wage reductions in a field dominated by a bourgeois political economy. It was perhaps as important for them to delegitimate factory owners' discursive constructions of the mill economy as it was to offer a viable alternative. Second, Owenism would be received sympathetically by other workers in the region but would have far less credibility among potential middle-class sympathizers who could exert pressure on the factory owners. Finally, one of the hallmarks of Owenism itself was that it offered a concept of production and distribution that provided an alternative to a wage-based economy. In an ironic sense, it offered the spinners only silence when it came to constructing their claims for a "just" wage.

The spinners instead seized upon the concepts of labor and property so central to political economy in constructing their repertoire. In so doing, they sought to legitimize their claims to what they perceived to be a "living wage" with the manufacturers' claims for profit. John Doherty, the secretary of the spinners' union, drew on the labor theory of value to assert the paramount sanctity of their piece-rate demands in the face of the manufacturers' intransigence during the strike of 1829: "Labour must give value to everything, and they who would reduce the price of labour were enemies of the country" (*SA*, Jan. 30, 1829). The Manchester spinners drew on the labor theory of value to disengage their claims for a just wage from the political economists' depictions of inexorable market fluctuations and reframed it as a matter of robbery. One of their weekly strike pamphlets noted,

> It is a simple principle of natural justice ... when one part of the community are [*sic*] wallowing in wealth, which they can scarcely consume or find use for, every other should be at least well fed, well clothed, and comfortably lodged. No class can better deserve these things, than those who produce *all* that is employed by the whole. The value of money may fluctuate. Its amount may be augmented or contracted at the will of a few. But these circumstances should have no influence on the condition of the labourer. There should be a fixed and understood quantity of food and clothing which every labourer should require for his family; and that whenever they come short of that quantity of food and clothing, they should look upon it, as so

> much taken from them, and never cease their exertions until it be restored to them again (PC, set 16, vol. 2, "Cotton," f. 72)

The labor theory of value was also turned into a vehicle for delegitimating the manufacturers' defense of their profits. In "The Cotton Spinners' and Power-Loom Weavers' Lesser Catechism," a biting commentary penned by an Ashton factory worker a few months prior to their great strike, the profit taking of manufacturers was depicted as an act of mammon worship and robbery. In this instance, the particular trope of the catechism exemplifies another facet of this dialogic process since, as I have noted, dissenting theology was part of the hegemonic discursive field.[33]

> Q. What does thou chiefly learn in these articles of belief?
> A. First, I learn to believe in the power of wealth, which giveth me all that I covet and not so to the rest of the world; secondly, in the accumulation of wealth by the application of other people's industry for my own ends. (*UTCJ*, May 22, 1830)

The spinners used these appropriations and redefinitions of the labor theory of value and the concept of property to develop another facet of their repertoire, their rights claims as productive members of the polity and society. With the manufacturers having defined the defense of property as part of the public good, working people could seize upon this argument and make a parallel case for their wage claims. In part, this discursive appropriation transformed these claims from private and individual matters conducted in the market to collective and public concerns. Doherty countered the attacks of the *Manchester Guardian* and *Stockport Advertiser* on the Ashton union and its strike through such an analogy: "They have been turned away for simply doing what the masters say they are doing, namely, protecting their property. The masters attempt, by a formidable combination, to take from the men a portion of that property they were then in possession of" (*Voice of the People,* March 5, 1831).

The spinners linked these rights claims to another facet of their reper-

---

[33] In the Ashton region, the factory owners' efforts for the betterment of working people focused largely on the organizing and underwriting of a number of Sunday schools (Rose 1969, pp. 12, 14–15; Hill 1902, pp. 110, 115–16). For many observers, the rationale behind this support was transparent. As Thomas Daniel, a Manchester mill owner, remarked, "I think the instructions given at those Sunday-schools are for the very purpose of making those children as humble and as obedient to the wishes of the manufacturers as possible" (BPP 1831–32, vol. XV, p. 327). By the mid-1820s, there were over 1,900 such child scholars in the Stalybridge and 4,100 in Ashton (Baines 1824, p. 556, Butterworth 1841, p. 104). By the early 1830s, from half to over three-quarters of all Ashton children were exposed to this Sunday edification, and it became almost a communal rite of maturation for working-class youth (BPP 1835, vol. XLI, pp. 80, 87, 422). The use of the trope itself can thus be seen as part of the counterhegemonic nature of the discursive repertoire.

toire, their discursive construction as productive citizens of a cherished nation. In their version of imagined community, workers were the bedrock of the citizenry, since as a banner carried during many of the Ashton strike parades proudly inveighed, "The labour of the nation is the wealth of the nation" (*Manchester Mercury* [hereafter *MM*], Dec. 21, 1830; *WMC*, Dec. 18, 1830). This parallels Doherty's nationalistic allusion to "enemies of the country" above and demonstrates the ways in which discourses concerning production could be linked to those of patriotism within the contemporary field.[34] The early 19th century was a period in which the constructions of national identity and patriotism were hotly contested, and even the image of the loyal plebeian "John Bull" was subject to spirited symbolic struggle (Colley 1992; Cunningham 1981; Taylor 1992). Asserting their status as the productive bedrock of the polity, they portrayed themselves as "true Britons," deserving of government protection from oppression.[35] The Manchester spinners even made a homological link between the functions of the state and that of the union (or "combination" in contemporary parlance), equating the purposes of governance in the two cases: "Government and Law is, or ought to be founded upon the PRINCIPLES of defending the weak and honest, against the powerful and unjust; and every well regulated community is a Combination for that purpose" (PC, set 16, vol. 2, "Cotton," f. 51). In doing so, emphases on individuality, negative rights, and minimal governance become displaced by those of collectivity, positive rights, and state activism for the communal good.

Two final pieces of the spinners' repertoire, their metaphorical and analogical incorporation of the discourse of the abolitionist movement and of radical republican politics, exemplify how the social and semiotic dimensions of possibilities within the discursive field shape the ways in which

---

[34] As I have argued in the case of the Spitalfields silk weavers for the same period, the discourse of citizenship was part of the discursive field that could be drawn upon to assert rights claims in production (Steinberg 1995*b*).

[35] In a similar vein, the Stockport spinners, commenting on the assistance of the nearby Hyde factory owners to their Stockport brethren by docking the wages of their workers who contributed to the Stockport strike fund, proclaimed, "Why then should they come forward, with a tyrannical impudence, never assumed by the worst despotism, and tell the free people of England, now working at Hyde, that they will fine them?" (PC, set 16, vol 2, f. 48). During this period, such discussion of the "free people of England" was part of a contested discourse of English Constitutionalism that was central to political debates, and can be seen as contiguous in the discursive field to constructions of "freedoms" in the market (Belchem 1981, 1988; Epstein 1994). Manufacturers of the region often emphasized the need for their own town M.P.'s through such nationalist discourses as well; for they proclaimed themselves as linchpins in the construction of Britannia around the world because of the role of cotton exports in creating British global markets

challengers can make claims. In each case, we see the complexity of the social, semiotic, and strategic relationality of discursive contention.

While England had itself abolished the slave trade in 1807, there were periodic campaigns within the country for a transatlantic abolition movement in the first decades of the 19th century. While drawing on a broad base of popular support, working-class radicals and factory reform advocates increasingly were agitated at what they perceived as the duplicity of the liberal monied ranks who ardently supported the cause of slave emancipation but generally remained cool, at best, to campaigns for working-class suffrage and factory reform legislation. Indeed, some vocal critics of factory reform in parliament were also champions of antislavery legislation (Drescher 1987, pp. 144–54; Hurwitz 1973, p. 43; Hollis 1980). As all three movements mounted petition campaigns and mass meetings in 1830–31, during the apex of the spinners' strikes, working-class radicals and union activists were quick to turn the discourse of antislavery against its liberal bourgeois advocates, among them Charles Hindley, as represented in the quote above. As David Turley observes of the discourse of abolitionism after the Napoleonic Wars, "Especially important were the growing certainties of liberal political economy and the theme of antislavery as expressing the national interest and, as an important moral component, the sense of national duty the British people owed the world and themselves" (1991, p. 18). Abolitionist discourse was thus constructed within a discursive field tied to popular political economy and the developing nationalism just noted.

With manufacturers such as Hindley and many prognosticators of political economy being vocal advocates of abolition, there was potential to appropriate a discourse of their exploiters, which highlighted issues of freedom and justice for laborers.[36] Factory workers thus often compared themselves and the slaves, and the "freedom" within abolitionist discourse was used to displace the "freedom" of the market in political economy. Ashton spinners' leader, J. J. Betts, argued that the slave had much greater freedom from want than the factory worker: "The only real difference is, that the negroes are slaves in name, while hundreds of thousands of our poor countrymen, here, are slaves in reality. There the slaves are comfortably housed, wholesomely fed, worked to the best economy of their health and strength, and I dare say sometimes overworked. Here,

---

[36] As David Turley observes of the abolitionist discourse and its uses among radicals and working-class critics of the factory, "In the case of radical critics of 'factory slavery' it was also a language sometimes used *against* abolitionists, illustrating the expansive and protean character of such language" (1991, p. 184). Abolitionist discourse was also strongly tied to evangelicalism and increasingly to nationalism during the 1820s and 1830s (Hurwitz 1973, pp. 40–43).

the slaves are miserably lodged, starved, beggarded, abused, despised, neglected, and overworked, always, and at all times without pity, without mercy, without hope" (*UTCJ,* May 8, 1830, p. 78).

Playing on the idea of the importance of fixed capital, a worker ingeniously explained why the slaves were in a superior position. The "free" labor of the market was transformed into a system of enslavement. Here again, we also see how the discourse of citizenship is also turned against the manufacturers: "The proprietor of the slave has an interest in his welfare. The return required on the capital sunk in the purchase, will induce him to feed and clothe him in such a way as to get the greatest amount of profit from his labour. . . . But there is no such motive to influence the conduct of the British capitalist. Those who employ thousands of 'freeborn' British artizans have no interest in their welfare beyond their labour of the day" (*UTCJ,* Apr. 3, 1830, p. 98).

In the late 1820s and early 1830s, England was rife with agitation for radical political reform, a cycle of protest that Tilly (1982, 1995*b*) depicts as the rise of the first social movement. Among working people, and especially factory workers in the North, there was broad support for radical political reform of Parliament and a deep-rooted cultivation of the discourses through which such claims were made. Factory owners were also champions of reforming Parliament, as many of their towns lacked direct representation. More generally, they chafed at their lack of influence relative to what they perceived as the rotted social and political foundations of aristocratic control, and as advocates of liberal reform, they railed against the illegitimate control of the lords. But there was often considerable dissension within these manufacturing districts, between manufacturers and factory workers, over the extent and pace of reform, with the former often arguing for property qualifications and the latter advocating universal manhood suffrage.

This marked political contention between factory owners and many of their workers created an opportunity to link political genres to industrial contention. Spinners' leaders, particularly in Ashton, creatively drew upon the discourse of radical republican politics to metaphorically justify their contentious (and sometimes violent) actions, a discourse that was part of the local established contention with manufacturers over the limits of parliamentary reform. The district was well known for the wide working-class support of radical politics (Cotton 1977, pp. 107, 109–11, 117–21, 127, 138, 141, 143; Hall 1989, p. 437; *MM,* Nov. 9, 30, 1830; *MT,* June 20, 1829, Nov. 27, 1830, Jan. 29, 1831; Lancashire County Record Office, DDX/880/2, Bowring to Lamb, Dec. 30, 1830; HO, ser. 40, box 26, ff. 46–50, Shaw to Bouverie, Aug. 29, 1830). The Ashton spinners' strike at the end of 1830 occurred not only at a time of tremendous agitation for parliamentary reform, but also concurrently with recent well-publicized

revolts against the monarchies in Poland and, more emblematically, France. The admonitions of two spinners' leaders at a large preparatory demonstration outside of Ashton provide an exemplary case of the ways that both contentious social relations and the boundaries set by discursive fields determined the ways in which actions could be framed. At this event, the leaders, Betts and Buckley, constructed the logic of righteous contention against the "Cotton Lords" through metaphorical appropriation and analogy, as the overthrow of aristocratic and theocratic tyranny provided the logic for the overthrow of economic tyranny:

> Betts observed that Trades Unions and Political Ones were now so intimately blended together that they must be looked upon as one. He proceeded to state that we lived under the Worst, the most Rascally, Despotic, Tyrannical Government that ever existed. He told the meeting of the Glorious Victory that had been achieved in France by only 8,000 Men over Tyranny and said there were more than 80,000 men ready for a similar proceeding in England. He then sat down apparently exhausted by the Efforts he had made and was followed by a person of the name of Buckley who resides in North Street near this place. . . . He stated that every Master was a Tyrant, that they had a right to participate in whatever property any Man had, that they must down with the Cotton Lords who had no right to any such profits, that they were Omnipotent in power, that if they would be United no Force could stand against them and that they must repel Force by force; they must rouse from their Apathy and let their Despotic Tyrannical Masters know theirs was the power and that they would use it. Betts again addressed them assuring them he fully concurred in the Sentiments of the last Speaker, he told them that this and other Meetings in the Villages were only preparatory to the great Meeting which would be held shortly in Staley Bridge or Ashton. He told them to recollect their Power was Omnipotent that they must shortly use it. . . He hoped that they would be united, that they would be determined to be Free. Let Liberty or Death be their Cry and Spreading out a small Flag with various Devices on it, told them that that was the Tricoloured Flag under which they must Act, that they must be Firm or this Opportunity would be Lost, and concluded with hoping that they would be united and All attend the meeting at Staley Bridge or Ashton as placards were to inform them, that all the Factories would Stop on that Day, and on that Day he intimated a Decisive Step would be taken. (HO, ser. 40, box 27, ff. 338–39, Nightingale to the Home Office, Nov. 8, 1830)[37]

During this turbulent time in November, the rebellions on the continent and domestic political agitation provided the spinners with analogies for

---

[37] The summary of this speech is a good example of the multivocality of discourse and its varying interpretations when it reaches actors to whom it is not intentionally directed. Nightingale, the author of the letter to the Home Office, was a minor excise officer who was traveling through the area and happened on the large open-air meeting. Not being fully apprised of the circumstances, he anxiously wrote a summary of the speeches to warn government officials that workers in the district were plotting revolution.

and exemplars of success. They reasoned their contest was paralleled to these revolutions, and the 80,000 members of their general union of all trades were spiritual kin to their rebellious brothers and sisters in France. The French tricolor was to be a standard symbol of the local unionists throughout the ensuing bitter strike, and the alarm of local authorities and the regional army commanders was piqued by its ubiquity. The tyranny of the "Cotton Lords" (as manufacturers were frequently termed) must be defeated just as that of the aristocracy. The relationality of the participants in contention and the available discourses within the discursive field allowed for a fertile hybridization of the critique of power. Local contention between liberal factory owners and working-class radicals over the pace and extent of political reform was mapped onto this economic conflict, which contained parallel fissures and opposing camps.

The spinners ultimately lost this and all but a few of the strikes to defend against piece-rate reductions. In the end, factory owners could afford to outlast their feisty underlings. Their defeats are a cogent reminder that success in collective action frequently hinges on material resources and access to the legitimate forces of repression, and that these two talk louder than words in the battle to determine who has the final say.

DISCUSSION

The spinners' discursive repertoire represents several facets of the dialogic perspective. First, we have seen that it was a relational social, semiotic, and strategic product. Their choice of representations and their attempts to convey a sense of injustice were significantly structured both by who their adversaries were (and their relations with them outside of this specific contention) and the structure of the discursive field dominated by these manufacturers. Though the spinners lived in well-knit and cohesive communities known for their distinctive working-class cultures, the discursive repertoire they fashioned was not an independent product created largely within their own subcultural networks. This is most apparent in their efforts to construct concepts of economic injustice. Rather than deploying a distinct economic discourse against political economy, or employing an alternative such as cooperationism, the spinners struggled within this hegemonic genre to establish the legitimacy of their claims. Contention was thus not between predominantly two distinct discourses or frames, but rather within a discursive field largely not of their own choosing.

Second, and relatedly, the spinners' construction of injustice and its resolution was not anchored in a single underlying collective action or master frame. Instead, the spinners produced a discursive repertoire within the discursive field largely defined by their employers and the au-

thorities and pundits who supported them. This repertoire was dominated by a selective appropriation of political economy, political liberalism, nationalism, abolitionism, and other genres through which factory owners mapped out a hegemonic vision of a social order. It also drew on the critique of power available from radical republican discourse in part because of the particular conjunction of the strike with contemporaneous political agitations. The repertoire was, following Franzosi, more of a "network of messages" whose coherence was developed over a series of contentious interactions with the factory owners and their allies (particularly the commercial newspaper editors).[38] In and of themselves, there was no necessary link for the spinners between the appropriations of the discourses of abolitionism and dissenting theology, and as I have noted, in other contexts, working-class radical activists indeed were cool to middle-class proselytization for both. Rather a complex flux of social, semiotic, and strategic relations determined how spinners fashioned their repertoire over time. Moreover, as I demonstrate elsewhere (Steinberg 1999), the predominance of all of the discourses in the spinners' repertoire ebbs and flows with the course of contention and related events in the political environment.

Third, the spinners' case demonstrates the multivocal nature of discourse and the ways in which the development of a repertoire is significantly an exercise in counterhegemony. The processes of appropriation through metaphor and analogy in which the spinners engaged highlight ways in which challengers, faced with a discursive field significantly not of their own choosing, sought to subvert powerholders' discursive dominance by exposing the interest-laden nature of their words. "Property," "rights," "freedom," "slave," and other key signifiers passed back and forth between the mouths of workers and employers, and in transformative processes of meaning.

Finally, the case of the spinners illuminates why it is problematic to draw analogies between discourse and material resources. Throughout their struggles, the spinners had no claims to exclusivity over the words they used to express injustice, nor did they exercise substantive control over them. Indeed, discursive conflict was dialectical, for in the process of making claims, the spinners also reinscribed themselves as partly captive to the truths of their opponents. In appropriating the discourse of political economy, for example, the spinners valorized their status as producers but did so by accepting the broad contours of political economy that legitimized a status for capitalists in the production process and the

---

[38] In Steinberg (1999), I demonstrate more extensively how this repertoire shifted over the course of these successive contests.

polity. The words they used were also prey to reappropriation by factory owners and hostile commentators, who could and did turn them back around against their working-class antagonists. The appropriation and belittling of a discourse of Christian piety also served to mark the spinners as suspicious atheistic radicals in the eyes of middle-class observers who might otherwise be sympathetic to their cause. Moreover, the use of the discourse of radical politics as an analogy for their contention was often misconstrued by local and national authorities as marking the spinners and their advocates as harbingers of political revolution (Steinberg 1999, chap. 9).

In addition, as I have noted above, some within the community who stood silently behind these words, particularly women factory workers, had their subordination *re*produced in other respects. The male spinners' interpretive repertoire reaffirmed the gendered hierarchy of the factory and community. Within its contemporary construction, the discourse of political economy structured a masculine coding of production and property. As Anna Clark (1992, 1995) notes in her analyses of working-class radical activism in the 1830s and 1840s, the vital support provided by women was often given at the expense of being subordinated through discourses that both explicitly and implicitly reinscribed their subordination. Indeed, as a variety of studies have recently shown, the explicit use of the feminine in factory agitation and labor struggle was keyed to a distinctly subordinate understanding of the role of women in the workplace and the economy (Gray 1996; Kodtischek 1997; Malone 1998; Rose 1992; Valenze 1995; Valverde 1988). For the female supporters of the spinners, therefore, the latters' struggles offered a distinctly mixed blessing.

CONCLUSION

Frame theory and its constructionist approach to contention has opened up the study of collective action to cultural analysis. In doing so, it has problemized the concept of interests and identity, broadened our understanding of mobilization, and brought the talk and back talk of contention centrally into our analyses. As I have suggested, however, despite such achievements, it only takes us part way to understanding the interactive relationships between social structures, cultural processes, and situational conjunctures. A dialogic analysis of interpretive repertoires surely does not provide us with all the answers, but I think it points the way to a more encompassing approach. In general, the spinners' case illustrates the advantages of this approach to cultural contention. The analysis of their discursive repertoire shows how such contention is keyed both to material situations and to structures of cultural meanings and codes that bound it.

Discursive repertoires are strategic, but collective actors are partly captives within the discursive fields that they seek to manipulate. Cultural codes do have a logic within which collective actors struggle, but accessing and transforming these codes is affected by structures and events beyond the noosphere of information. Only by understanding the causal interactions between social structures, situated action, and culture can we fully appreciate the way any one of this triad bounds the course of contention.

In this approach, we should analyze the duality of culture in contention, the ways in which it serves as a map for struggle, but also, as Fantasia and Hirsch (1995) have argued, as a contested terrain as well. This means, as Johnston and Klandermans (1995) suggest, that we must move beyond the analysis of social movement culture as reified codes or texts. Only when we see cultural processes as part of the action, as part of what is quintessential both in formation and at stake, can we fully appreciate the cultural dimension of collective action. To do so, as I have argued, we have to deepen our theoretical perspective on the importance of talk and back talk in contentious action. Dialogism offers a specific framework for a more dynamic analysis of collective action discourse contextually keyed to ongoing hegemonic struggle. It focuses attention on the discursive repertoires produced by challengers, how these repertoires often are fashioned through an ongoing interaction with the powerholders' genres, and the continual uncertainties and challenges that these repertoires pose for all involved.

The back talk of power's discontents often is an effort to refigure truth, redefine justice, and usurp powerholders' own moral authority by snatching their words from their mouths. "There is neither a first nor a last word," noted Bakhtin (1986, p. 166). Dialogic analysis shows how this insight informs our understanding of contention, both in terms of how the powerful use the word to create truth, and how challengers reach within the word to turn it around toward a better, if uncertain, future.

REFERENCES

Baines, Edward. 1824. *History, Directory and Gazetteer of the County Palantine of Lancashire,* vol. 2. Liverpool: Wm. Wales.
————. (1835) 1966. *History of the Cotton Manufacture in Great Britain.* London: Frank Cass.
Bakhtin, Mikhail M. 1986. *Speech Genres and Other Late Essays,* translated by Vern W. McGee. Edited by Caryl Emerson and Michael Holquist. Austin: University of Texas Press.
————. 1984 *Problems of Dostoevsky's Poetics,* edited and translated by Caryl Emerson. Minneapolis: University of Minnesota Press.
————. 1981. *The Dialogic Imagination: Four Essays,* translated by Caryl Emerson and Michael Holquist. Edited by Michael Holquist. Austin: University of Texas Press.

Barker, Colin. 1997. "Social Confrontation in Manchester's Quangoland: Local Protest over the Proposed Closure of Booth Hall Children's Hospital." *North West Geographer*, 3d series, 1 (1): 18–28.

Belchem, John. 1981 "Republicanism, Popular Constitutionalism and the Radical Platform in Nineteenth-Century England." *Social History* 6:1–32.

———. 1988. "Radical Language and Ideology in Nineteenth-Century England: The Challenge of the Mass Platform." *Albion* 20·247–59.

Bell, Michael Mayerfield. 1998. "Culture as Dialogue." Pp. 49–62 in *Bakhtin and the Human Sciences*, edited by Michael Mayerfield Bell and Michael Gardiner. London: Sage.

Benford, Robert. 1997. "An Insider's Critique of the Social Movement Framing Perspective " *Sociological Inquiry* 67 (4): 409–30.

Berg, Maxine. 1980 *The Machinery Question and the Making of Political Economy*. Cambridge: Cambridge University Press.

Bourdieu, Pierre, and Loic J. D Wacquant 1992. *An Invitation to Reflexive Sociology*. Chicago. University of Chicago Press

BPP (British Parliamentary Paper). 1831–32 (706) XV. Report of the Select Committee on the Bill to Regulate the Labor of Children in Mills and Factories.

———. 1833 (690) VI. Report of the Select Committee on the Present State of Manufactures, Commerce, and Shipping in the United Kingdom.

———. 1837–38 (485) VIII. First Report from the Select Committee on Combinations of Workmen.

Brandist, Craig. 1996a. "Gramsci, Bakhtin and the Semiotics of Hegemony." *New Left Review*, 246·94–109.

———. 1996b. "The Official and the Popular in Gramsci and Bakhtin." *Theory Culture and Society* 13 (2): 59–74

Burkitt, Ian. 1998. "The Death and Rebirth of the Author· The Bakhtin Circle and Bourdieu on Individuality, Language and Representation." Pp. 163–80 in *Bakhtin and the Human Sciences*, edited by Michael Mayerfield Bell and Michael Gardiner. London: Sage.

Butterworth, Edwin. 1841. *A Statistical Sketch of the County Palatine of Lancaster*. London. Longman.

Capek, Stella M. 1993. "The 'Environmental Justice' Frame: A Conceptual Discussion and an Application." *Social Problems* 40:5–24.

Carroll, William K., and R. S. Ratner. 1994. "Between Leninism and Radical Pluralism: Gramscian Reflections on Counter-Hegemony and the New Social Movements." *Critical Sociology* 20 (2)· 2–26.

———. 1996. "Master Framing and Cross-Movement Networking in Contemporary Social Movements." *Sociological Quarterly* 37 (4): 601–25

Catling, Harold. 1970. *The Spinning Mule* Newton Abbot: David & Charles.

Claeys, Gregory. 1985. "The Reaction to Political Radicalism and the Popularisation of Political Economy in Early Nineteenth-Century Britain." Pp. 119–36 in *Expository Science: Forms and Functions of Popularisation*, edited by Terry Shinn and Richard Whitley. Dordrecht: D. Reidel.

———. 1987. *Machinery, Money, and the Millennium: From Moral Economy to Socialism, 1815–1850*. Cambridge: Polity

———. 1989. *Citizens and Saints: Politics and Anti-Politics in Early British Socialism*. Cambridge: Cambridge University Press

Clark, Anna. 1992. "The Rhetoric of Chartist Domesticity· Gender, Language, and Class in the 1830s and 1840s " *Journal of British Studies* 31:62–88.

———. 1995. *The Struggle for the Breeches: The Making of the British Working Class, 1780–1850*. Berkeley and Los Angeles: University of California Press.

Clark, Katrina, and Michael Holquist. 1984. *Mikhail Bakhtin*. Cambridge, Mass.: Harvard University Press.

Colley, Linda. 1992. *Britons: Forging the Nation 1707–1837.* New Haven, Conn.: Yale University Press.

Collins, Chik. 1999. *Language, Ideology and Social Consciousness: Developing a Sociohistorical Approach.* Aldershot: Ashgate.

Cotton, Nicolas. 1977. Popular Movements in Ashton-under-Lyne and Stalybridge before 1832. M.Litt. Thesis. University of Birmingham, Department of History.

Crossley, Nick. 1999. "Habitus, Capital and Repertoires of Contention in Recent Mental Health Processes." Paper presented at the Fifth International Conference on Alternative Futures and Popular Protest, Manchester Metropolitan University, March 31.

Cunningham, Hugh. 1981. "The Language of Patriotism, 1750–1914." *History Workshop* 12:8–37.

Doherty, John. 1829. *A Report of the Proceedings of a Delegate Meeting, of the Operative Spinners of England, Ireland and Scotland, Assembled at Ramsey, Isle of Man, On Saturday, December 5, 1829, and Three Following Days.* Manchester: M. Wardle.

Donati, Paolo R. 1992. "Political Discourse Analysis." Pp. 136–67 in *Studying Collective Action,* edited by Mario Diani and Ron Eyerman. London: Sage

Drescher, Seymour. 1987. *Capitalism and Antislavery: British Mobilization in Comparative Perspective.* Oxford: Oxford University Press.

Ellingson, Stephen J. 1995. "Understanding the Dialectic of Discourse and Collective Action: Public Debate and Rioting in Antebellum Cincinnati." *American Journal of Sociology* 101:100–44.

Epstein, James. 1994. *Radical Expression: Political Language, Ritual, and Symbol in England, 1790–1850.* New York: Oxford University Press.

Evans, John. 1997. "Multi-Organizational Fields and Social Movement Organization Frame Content: The Religious Pro-Choice Movement." *Sociological Inquiry* 67 (4): 451–70.

Ewick, Patricia, and Susan Silbey. 1995. "Subversive Stories and Hegemonic Tales: Toward a Sociology of Narrative." *Law and Society Review* 29 (2): 197–226.

Fantasia, Rick, and Eric L. Hirsch. 1995. "Culture in Rebellion: The Appropriation and Transformation of the Veil in the Algerian Revolution." Pp. 144–59 in *Social Movements and Culture,* edited by Hank Johnston and Bert Klandermans. Minneapolis: University of Minnesota Press.

Fetter, Frank W 1965. "Economic Controversy in British Reviews, 1802–1850." *Economica,* new series, 32 (128): 426–37

Fine, Gary Alan. 1995. "Public Narration and Group Culture: Discerning Discourse in Social Movements." Pp. 127–43 in *Social Movements and Culture,* edited by Hank Johnston and Bert Klandermans. Minneapolis. University of Minnesota Press.

Follows, John W. 1951. *Antecedents of the International Labor Organization.* Oxford: Clarendon Press.

Fowler, Alan, and Terry Wyke, eds. 1987 *The Barefoot Aristocrats: A History of the Amalgamated Association of Operative Cotton Spinners.* Littleborough: George Kesall.

Franzosi, Roberto. 1998. "Narrative Analysis—or Why (and How) Sociologists Should Be Interested in Narrative." *Annual Review of Sociology* 24:505–42.

———. In press. *From Words to Numbers.* Cambridge: Cambridge University Press.

Freifeld, Mary. 1986. "Technological Change and the Self-Acting Mule: A Study of Skill and the Sexual Division of Labor." *Social History* 11:319–43.

Gamson, Joshua. 1995. "Must Identity Movements Self-Destruct? A Queer Dilemma." *Social Problems* 42:390–407.

Gamson, William A. 1988*a* "Political Discourse and Collective Action." *International Social Movement Research* 1:219–44.

——— 1988b. "The 1987 Distinguished Lecture: A Constructionist Approach to Mass Media and Public Opinion." *Symbolic Interaction* 11:161–74.

———. 1992a. "The Social Psychology of Collective Action." Pp. 53–76 in *Frontiers in Social Movement Theory*, edited by Aldon D Morris and Carol McClung Mueller. New Haven, Conn.: Yale University Press.

———. 1992b. *Talking Politics* Cambridge: Cambridge University Press.

———. 1995. "Constructing Social Protest." Pp. 85–106 in *Social Movements and Culture*, edited by Hank Johnston and Bert Klandermans Minneapolis: University of Minnesota Press.

Gamson, William A., Bruce Fireman, and Steven Rytina. 1982. *Encounters with Unjust Authority*. Homewood, Ill Dorsey Press.

Gamson, William A., and David S. Meyer. 1996. "Framing Political Opportunity." Pp. 275–90 in *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures and Cultural Framings*, edited by Doug McAdam, John D. McCarthy, and Mayer Zald. Cambridge: Cambridge University Press.

Gamson, William A., and Andre Modigliani. 1989. "Media Discourse and Public Opinion on Nuclear Power: A Constructionist Approach." *American Journal of Sociology* 95:1–37.

Gardiner, Michael. 1992. *The Dialogics of Critique: M. M. Bakhtin and the Theory of Ideology*. London: Routledge.

Gerhards, Jürgen, and Dieter Rucht. 1992. "Mesomobilization: Organizing and Framing in Two Protest Campaigns in West Germany." *American Journal of Sociology* 98:555–95.

Goffman, Erving. 1974. *Frame Analysis: An Essay on the Organization of Experience.* Cambridge, Mass.: Harvard University Press.

———. 1981. *Forms of Talk*. Philadelphia: University of Pennsylvania Press.

Goldstrom, Max. 1985. "Popular Political Economy for the British Working Class in the Nineteenth Century " Pp. 259–73 in *Expository Science: Forms and Functions of Popularisation*, edited by Terry Shinn and Richard Whitley. Dordrecht: D. Reidel.

Gottdiener, Mark 1995. *Postmodern Semiotics. Material Culture and the Forms of Postmodern Life*. Oxford: Blackwell.

Gray, Robert. 1996. *The Factory Question and Industrial England, 1830–1860*. Cambridge: Cambridge University Press.

Hall, Catherine. 1992 *White, Male and Middle Class: Explorations in Feminism and History*. London: Routledge.

Hall, Robert G. 1989. "Tyranny, Work and Politics: The 1818 Strike Wave in the English Cotton District." *International Review of Social History* 34:433–70.

———. 1991. *Work, Class and Politics in Ashton-under-Lyne, 1830–1860*. Ph.D. dissertation. Vanderbilt University, Department of History.

Hart, Stephen. 1996 "The Cultural Dimension of Social Movements: A Theoretical Reassessment and Literature Review." *Sociology of Religion* 57:87–100.

Herndl, Diane. 1991. "The Dilemmas of a Feminine Dialogic." Pp. 7–24 in *Feminism, Bakhtin and the Dialogic*, edited by Dale M. Bauer and Susan J. McKinstry. Albany: State University of New York Press.

Hill, Samuel. 1902. *Bygone Stalybridge*. Private printing.

Hindley, Charles 1841. *Free Trade. Speech of Charles Hindley, Esq. Now M. P. for Ashton-under-Lyne, at the Public Dinner Given in Saddleworth, to Lord Morpeth and Henry Brougham Esq., Sept. 25, 1830. On Their Election for the County of York*. London: James Ridgway.

Hirschman, Albert O. 1986. *Rival Views of Market Societies and Other Recent Essays*. New York: Penguin.

Hitchcock, Peter. 1994. *Dialogics of the Oppressed.* Minneapolis. University of Minnesota Press.

HO (British Public Record Office, Home Office Papers). Ser. 40, boxes 26–27. Correspondence, Civil Disturbances.

Hodge, Robert, and Gunther Kress. 1988. *Social Semiotics*. Ithaca, N.Y · Cornell University Press.

Hollis, Patricia. 1980. "Anti-Slavery and British Working-Class Radicalism in the Years of Reform." Pp. 294–315 in *Anti-Slavery, Religion and Reform: Essays in Memory of Roger Anstey*. Folkestone. William Dawson & Sons.

Holquist, Michael. 1990. *Dialogism: Bakhtin and His World*. London: Routledge.

Hopkins, Thomas. 1831. *Wages, or Masters and Workmen*. Manchester: Alexander Wilson.

Horne, Thomas A. 1990. *Property Rights and Poverty: Political Argument in Britain, 1605–1834*. Chapel Hill: University of North Carolina Press.

Howe, Anthony. 1984 *The Cotton Masters, 1830–1860*. Oxford: Clarendon Press.

Hoy, Mikita. 1992. "Bakhtin and Popular Culture." *New Literary History* 23 765–82.

Huberman, Michael. 1986. "Invisible Handshakes in Lancashire: Cotton Spinning in the First Half of the Nineteenth Century." *Journal of Economic History* 46 (4): 987–98.

Hunt, Alan. 1990. "Rights and Social Movements: Counter-Hegemonic Strategies." *Journal of Law and Society* 17 (3): 309–28.

Hunt, Scott A., Robert D. Benford, and David A. Snow. 1994. "Identity Fields: Framing Processes and the Social Construction of Movement Identities." Pp. 185–208 in *New Social Movements: From Ideology to Identity*, edited by Enrique Laraña, Hank Johnston, and Joseph R. Gusfield. Philadelphia: Temple University Press

Hurwitz, Edith F. 1973. *Politics and the Public Conscience: Slave Emancipation and the Abolitionist Movement in Britain*. London: George Allen & Unwin.

Jasper, James M. 1997. *The Art of Moral Protest: Culture, Biography and Creativity in Social Movements*. Chicago: University of Chicago Press.

Johnston, Hank. 1991. *Tales of Nationalism: Catalonia, 1939–1979*. New Brunswick, N.J.: Rutgers University Press.

———. 1995. "A Methodology for Frame Analysis: From Discourse to Cognitive Schemata." Pp. 217–46 in *Social Movements and Culture*, edited by Hank Johnston and Bert Klandermans Minneapolis: University of Minnesota Press.

Johnston, Hank, and Bert Klandermans. 1995 "The Cultural Analysis of Social Movements." Pp 3–24 in *Social Movements and Culture*, edited by Hank Johnston and Bert Klandermans. Minneapolis: University of Minnesota Press.

Kane, Anne E. 1997. "Theorizing Meaning Construction in Social Movements: Symbolic Structures and Interpretation during the Irish Land War, 1879–1882." *Sociological Theory* 15 (3): 249–76.

Kirby, R. G., and A. E. Musson. 1975. *The Voice of the People: John Doherty, 1798–1854*. Manchester: Manchester University Press.

Klandermans, Bert. 1988. "The Formation and Mobilization of Consensus." *International Social Movement Research* 1:173–96.

———. 1992. "The Social Construction of Protest and Multiorganizational Fields." Pp. 77–103 in *Frontiers in Social Movement Theory*, edited by Aldon D. Morris and Carol McClung Mueller. New Haven, Conn.: Yale University Press.

———. 1997. *The Social Psychology of Protest*. Oxford: Blackwell.

Knight, Charles. 1831. *A Short Address to Workmen, on Combinations to Raise Wages*. London: W. Clowes.

Kniss, Fred. 1996. "Ideas and Symbols as Resources in Intrareligious Conflict: The Case of American Mennonites" *Sociology of Religion* 57:7–23

———. 1997. *Disquiet in the Land: Cultural Conflict in American Mennonite Communities*. New Brunswick, N.J.: Rutgers University Press.

Koditschek, Theodore. 1990. *Class Formation and Urban-Industrial Society: Bradford, 1750–1850*. Cambridge: Cambridge University Press.

————. 1997. "The Gendering of the British Working Class." *Gender and History* 9 (2): 333–63.

Lazonick, William. 1979. "Industrial Relations and Technical Change: The Case of the Self-Acting Mule." *Cambridge Journal of Economics* 3·231–62.

————. 1981. "Production Relations, Labor Productivity, and Choice of Technique: British and U.S. Cotton Spinning." *Journal of Economic History* 41.491–516.

Lilley, Stephen J., and Gerald M. Platt. 1994. "Correspondents' Images of Martin Luther King, Jr.: An Interpretive Theory of Movement Leadership." Pp. 65–83 in *Constructing the Social*, edited by Theodore R. Sarbin and John I. Kitsuse. London: Sage.

Macdonell, Diane. 1986. *Theories of Discourse: An Introduction*. Oxford: Blackwell.

Malone, Carolyn. 1998. "Gendered Discourses and the Making of Protective Labor Legislation in England, 1830–1914 " *Journal of British Studies* 37 (2): 166–91.

Manning, P. K., and Betsy Cullum-Swan. 1992 "Semiotics and Framing: Examples." *Semiotica* 92:239–57.

Marullo, Sam, Ron Pagnucco, and Jackie Smith 1996. "Frame Changes and Social Movement Contraction: U.S. Peace Movement Framing after the Cold War." *Sociological Inquiry* 66 (1): 1–28.

Masson, Dominique. 1996. "Language, Power and Politics: Revisiting the Symbolic Challenge of Movements." *Alternate Routes* 13:67–100.

McAdam, Doug. 1996. "The Framing Function of Movement Tactics: Strategic Dramaturgy in the American Civil Rights Movement." Pp. 338–55 in *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures and Cultural Framings*, edited by Doug McAdam, John D. McCarthy, and Mayer Zald. Cambridge: Cambridge University Press.

————. 1994. "Culture and Social Movements." Pp. 36–58 in *New Social Movements: From Ideology to Identity*, edited by Enrique Laraña, Hank Johnston, and Joseph R. Gusfield. Philadelphia: Temple University Press.

McAdam, Doug, John D. McCarthy, and Mayer Zald. 1996. "Introduction: Opportunities, Mobilizing Structures, and Framing Processes—toward a Synthetic, Comparative Perspective on Social Movements." Pp. 1–20 in *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures and Cultural Framings*, edited by Doug McAdam, John D. McCarthy, and Mayer Zald. Cambridge: Cambridge University Press.

McAdam, Doug, Sidney Tarrow, and Charles Tilly. 1997. "Toward an Integrated Perspective on Social Movements and Revolution." Pp. 142–73 in *Comparative Politics: Rationality, Culture and Structure*, edited by Mark I. Lichbach and Alan S. Zuckerman. Cambridge: Cambridge University Press.

McCarthy, John D., Jackie Smith, and Mayer N. Zald. 1996. "Accessing Public, Media, Electoral and Governmental Agendas." Pp. 291–311 in *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures and Cultural Framings*, edited by Doug McAdam, John D. McCarthy, and Mayer Zald. Cambridge: Cambridge University Press.

Medvedev, P. N., and M. M. Bakhtin. 1978. *The Formal Method in Literary Scholarship: A Critical Introduction to Sociological Poetics*, translated by Albert J. Wehrle. Baltimore: Johns Hopkins University Press.

Mills, Sara. 1997. *Discourse*. London: Routledge.

Mooney, Patrick H., and Scott A Hunt. 1996. "A Repertoire of Interpretations: Master Frames and Ideological Continuity in U.S. Agrarian Mobilization." *Sociological Quarterly* 37:177–97.

Morris, Robert J. 1990. *Class, Sect and Party: The Making of the Middle Class, Leeds 1820–1850*. Manchester· Manchester University Press.

Morson, Gary Saul, and Caryl Emerson. 1990 *Mikhail Bakhtin: Creation of a Prosaics*. Stanford, Calif.: Stanford University Press

Pateman, Carole. 1988. *The Sexual Contract.* Stanford, Calif.: Stanford University Press.

PC (Francis Place Collection of Newspaper Clippings and Pamphlets). British Library, London.

Platt, Gerald, and Michael Fraser. 1998. "Race and Gender Discourse Strategies: Creating Solidarity and Framing the Civil Rights Movement." *Social Problems* 45 (2): 160–79.

Platt, Gerald M., and Stephen J. Lilley. 1994. "Multiple Images of a Charismatic: Constructing Martin Luther King Jr.'s Leadership." Pp. 55–74 in *Self, Collective Behavior and Society: Essays in Honor of the Contributions of Ralph H. Turner,* edited by Gerald M. Platt and Chad Gordon. Greenwich, Conn.: JAI Press.

Polletta, Francesca. 1997 "Culture and Its Discontents: Recent Theorizing on the Cultural Dimensions of Protest." *Sociological Inquiry* 67 (4): 431–50.

———. 1998. "'It Was Like a Fever . . .' Narrative Identity in Social Protest." *Social Problems* 45 (2): 137–59.

Ponzio, Augusto. 1990 *Man as Sign: Essays on the Philosophy of Language,* edited and translated by Susan Petrilli. New York: Mouton de Gruyter.

Rose, A E 1969 *Methodism in Ashton-under-Lyne, Part II.* Ashton-under-Lyne: Andrews & Co.

Rose, Sonya O. 1992. *Limited Livelihoods: Gender and Class in Nineteenth-Century England.* Berkeley: University of California Press.

———. 1993. "Gender and Labor History: The Twentieth-Century Legacy." *International Review of Social History,* supp.:145–62.

Scott, James C. 1990. *Domination and the Arts of Resistance: Hidden Transcripts.* New Haven, Conn.: Yale University Press.

Seed, John. 1982. "Unitarianism, Political Economy, and the Antimonies of Liberal Culture in Manchester, 1830–50." *Social History* 7:1–25.

———. 1986 "Theologies of Power: Unitarianism and the Social Relations of Religious Discourse, 1800–50." Pp. 108–56 in *Class, Power and Social Structure in Nineteenth-Century Religious Towns,* edited by R. J. Morris. Leicester: Leicester University Press.

Shevstova, Maria. 1992. "Dialogism in the Novel and Bakhtin's Theory of Culture." *New Literary History* 23:747–63.

Shotter, John. 1992. "Bakhtin and Billig: Monological versus Dialogical Practices " *American Behavioral Scientist* 36:8–21.

Shotter, John, and Michael Billig. 1998. "A Bakhtinian Psychology: From out of the Heads of Individuals into the Dialogues between Them." Pp. 13–29 in *Bakhtin and the Human Sciences,* edited by Michael Mayerfield Bell and Michael Gardiner. London: Sage.

Snow, David A., and Robert D. Benford. 1988. "Ideology, Frame Resonance, and Participant Mobilization." *International Social Movement Research* 1:197–217.

———. 1992. "Master Frames and Cycles of Protest." Pp. 133–55 in *Frontiers in Social Movement Theory,* edited by Aldon D. Morris and Carol McClung Mueller. New Haven, Conn.: Yale University Press.

Snow, David A., E. Burke Rochford, Jr., Steven K. Worden, and Robert D. Benford. 1986. "Frame Alignment Processes, Micromobilization and Movement Participation." *American Sociological Review* 51:464–81.

Somers, Margaret. 1994. "The Narrative Construction of Identity: A Relational and Network Approach." *Theory and Society* 23 (5): 605–49.

Somers, Margaret, and Gloria Gibson 1994. "Reclaiming the Epistemological 'Other': Narrative and the Social Construction of Identity." Pp. 37–99 in *Social Theory and the Politics of Identity,* edited by Craig J. Calhoun. Oxford: Blackwell.

Spillman, Lyn. 1995. "Culture, Social Structures and Discursive Fields." *Current Perspectives in Social Theory* 15:129–54.

Steinberg, Marc W. 1995a. "Repertoires of Discourse: The Case of the Spitalfields Silk Weavers and the Moral Economy of Conflict." Pp 57–88 in *Cycles and Repertoires in Collective Action,* edited by Mark Traugott. Durham, N.C.: Duke University Press.

———. 1995b. " 'The Great End of All Government . . .': Working Peoples' Construction of Citizenship Claims in Early Nineteenth-Century England and the Matter of Class." *International Review of Social History* 40:19–50.

———. 1998. "Tilting the Frame: Considerations on Collective Action Framing from a Discursive Turn." *Theory and Society* 27 (6): 845–72.

———. 1999. *Fighting Words: Working-Class Formation, Collective Action and Discourse in Early Nineteenth-Century England* Ithaca, N.Y.: Cornell University Press.

Sykes, Robert A. 1982. "Popular Politics and Trade Unionism in Southeast Lancashire 1829–1842," vol. 1. Ph.D. dissertation University of Manchester, Department of History.

Tarrow, Sidney. 1994. *Power in Movement: Social Movements, Collective Action and Politics.* Cambridge: Cambridge University Press.

Taylor, Miles 1992. "John Bull and the Iconography of Public Opinion in England c. 1712–1929." *Past and Present* 134:93–128.

Taylor, Verta, and Nancy Whittier 1995. "Analytical Approaches to Social Movement Culture: The Culture of the Women's Movement." Pp. 163–87 in *Social Movements and Culture,* edited by Hank Johnston and Bert Klandermans Minneapolis: University of Minnesota Press

Thompson, Noel W. 1984. *The People's Science· The Popular Political Economy of Exploitation and Crisis 1816–34.* Cambridge Cambridge University Press.

———. 1988. *The Market and Its Critics: Socialist Political Economy in Nineteenth Century Britain.* London. Routledge.

Tilly, Charles. 1982 "Britain Creates the Social Movement " Pp. 21–51 in *Social Conflict and Political Order in Modern Britain,* edited by James Cronin and Jonathan Schneer. New Brunswick, N.J · Rutgers University Press.

———. 1995a. "Contentious Repertoires in Great Britain, 1758–1834." Pp 15–42 in *Repertoires and Cycles of Collective Action,* edited by Mark Traugott. Durham, N.C.: Duke University Press

———. 1995b. *Popular Contention in Great Britain, 1758–1834.* Cambridge, Mass.: Harvard University Press.

Todorov, Tzvetan. 1984. *Mikhail Bakhtin: The Dialogic Principle.* Minneapolis: University of Minnesota Press.

Tufnell, Edward C 1834. *Character, Object and Effects of Trades' Unions: With Some Remarks on the Law Concerning Them.* London: James Ridgway & Sons

Turley, David. 1991. *The Culture of English Antislavery, 1780–1860.* London: Routledge.

Turner, Herbert A. 1962. *Trade Union Growth, Structure and Policy: A Comparative Study of the Cotton Unions.* London: George Allen & Unwin.

Valenze, Deborah. 1995. *The First Industrial Woman.* Oxford: Oxford University Press.

Valverde, Marianna. 1988. "'Giving the Female a Domestic Turn': The Social, Legal and Moral Regulation of Women's Work in British Cotton Mills, 1820–1850." *Journal of Social History* 21:619–34.

Volosinov, V. N 1986. *Marxism and the Philosophy of Language.* Cambridge, Mass · Harvard University Press.

———. 1983. "Literary Stylistics." Pp. 93–152 in *Bakhtin School Papers,* edited by Ann Shukman. Oxford: RPT Publications.

Voss, Kim. 1993. *The Making of American Exceptionalism: The Knights of Labor and Class Formation in the Nineteenth Century.* Ithaca, N.Y · Cornell University Press.

Williams, Rhys H. 1995 "Constructing the Public Good: Social Movements and Cultural Resources." *Social Problems* 42:124–44

Zald, Mayer. 1996. "Culture, Ideology and Strategic Framing." Pp. 261–74 in *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures and Cultural Framings,* edited by Doug McAdam, John D. McCarthy, and Mayer Zald. Cambridge: Cambridge University Press.

# Conjuncture, Comparison, and Conditional Theory in Macrosocial Inquiry[1]

Jeffery M. Paige
*University of Michigan*

Recent debates on comparative historical method have focused on three issues: (1) narrative and conjuncture, (2) the logic of comparison, and (3) the limits of theoretical generalization. The present article attempts to resolve some of the issues raised in these debates by developing a distinction between universal and historically conditional theory through a contrast between work by such leaders of the revived comparative historical tradition as Charles Tilly, Theda Skocpol, and Immanuel Wallerstein and selected works by second-generation comparative historical sociologists. The conditional theories of the second generation incorporate narrative and conjunctural temporality, theory-driven comparison, and historically conditional generalization that were not emphasized in the universalizing theories and comparisons of the first generation.

The reemergence of comparative historical sociology in the 1970s and the adoption of this research style by many younger scholars in the 1980s and 1990s has stimulated debate concerning both its epistemological foundations and implications for general theory in the social sciences. The debate has increasingly organized itself around three issues: historical time, with its emphasis on sequence, event, conjuncture, and narrative; the logic of comparison, particularly among small numbers of large units defined specifically in time and space; and the problem of generalization and causal theory. Concern with historic specifics of time and space have led a number of the most thoughtful commentators (Abbott 1990, 1992*a*, 1992*b*; Griffin 1992, 1993; Quadagno and Knapp 1992; Sewell 1996; Somers 1996, 1998; Steinmetz 1998; Stinchcombe 1978; Tomich 1995) to ques-

tion the desirability or even the possibility of general theory in social re-search. This implication of comparative historical work has been vigorously resisted by others (Burawoy 1989; Kiser and Hechter 1991).

There is considerable agreement in the recent literature on historical time that, as Tilly (1994, p. 270) notes, for historians, "time and space are usually thick—drenched with causes that inhere in sequence, accumula-tion, contingency and proximity." The recent literature has emphasized two related aspects of historical time that have particular significance for sociological explanation—narrative and conjuncture. Narrative explana-tion views events themselves as causes and assumes that contingent se-quences of events limit future possibilities and determine action (Abbott 1990, 1992a, 1992b; Aminzade 1992; Griffin 1992, 1993; Somers 1996, 1998; Sewell 1996). Conjunctural explanation emphasizes that a particular combination of structural causes and events, in a particular time and place, may create unique outcomes that will not necessarily be repeated in other contexts (Griffin 1992; Ragin 1987; Sewell 1996; Steinmetz 1998; Tomich 1995).

While there is agreement about the causal importance of these aspects of historical time, there is less agreement about implications of such a conception of time for sociological generalization. The dominant view in recent work seems to be that, as Griffin (1992, p. 415) says, historical expla-nation "requires a logic of explanation quite different in form from that of subsumptive explanation, relying on . . . comparison, externalization and generalization." Sewell (1996, p. 272) concludes that "sociology's epic quest for social laws is illusory." Steinmetz (1998, pp. 184, 182 n. 28) argues that "contingent, conjunctural causality is the norm in open systems like society" requiring an "eclectic mix of theories." Somers (1998, p. 766; em-phasis in original) postulates "relational realism" in which "causal power . . . is independent from belief in any *one* particular theory." Abbott (1992b, p. 68) argues that "things happen because of constellations of factors, not because of a few fundamental factors acting independently." Stinchcombe (1978, p. 22) claims that good historical sociology depends on locating "deep historical analogies," which are largely independent of general causal theories. Skocpol (1979, p. 190) rejects the "blinders" of "pre-existing theories." Quadagno and Knapp (1992) attempt to demonstrate that causal analysis is possible without appeal to general theories.

On the other hand, Michael Burawoy (1989) and Edgar Kiser and Mi-chael Hechter (1991) claim that the goal of comparative historical inquiry is to advance general theories of society, not to explain specific historical conjunctures. Burawoy argues, after Lakatos, that scientific research is organized in research programs in which the implications of core theoreti-cal postulates are successfully reformulated into new "belts" of theory to account for "anomalies" or disconfirmations of predictions made by the

original theory. Kiser and Hechter deny that any research, including historical research, is possible without general theory and reject what they see as the historicism and inductivism of much comparative historical sociology.[2]

The debate about the role of theory in historical temporality is echoed in the debate stimulated by Theda Skocpol's *States and Social Revolutions: A Comparative Analysis of France, Russia and China* (1979) over the role of case logic and Millian methods of systematic comparison in historical analysis. Kiser and Hechter (1991) and Burawoy (1989) argue that Skocpol's inductivism downplays the role of theory in comparative historical research, while Lieberson (1992), Ragin (1987), and Nichols (1986) have questioned the validity of such induction from small numbers of historical cases. Sewell (1996) and Abbott (1992a), on the other hand, have criticized Skocpol for a neglect of historical narrative and conjuncture in her case comparisons. Once again, much of the debate revolves around the question of generality from the historically specific—this time as much in space as in time.

I hope to make a contribution to these debates by contrasting the work of three leaders of the seventies-generation revival—Charles Tilly, Theda Skocpol, and Immanuel Wallerstein—with three works of second-generation scholars who, it will be argued, are solving in practice the issues of temporality, comparison, and generalization posed in recent debate. Howard Kimeldorf's *Reds or Rackets: The Making of Radical and Conservative Unions on the Waterfront* (1988) is centrally concerned with questions of historical narrative and conjuncture. William Sewell (1996) has contrasted Kimeldorf's historically "eventful" conception of time with the "abstract transhistorical time" of Tilly and Wallerstein. Gay Seidman's *Manufacturing Militance: Workers' Movements in Brazil and South Africa 1970–1985* (1994) like Theda Skocpol's *States and Social Revolutions* is a work of cross-societal comparison based on Mill's "method of similarity." Seidman, however, does not proceed inductively in the manner of Skocpol, but rather uses her comparisons to search for observations that are theoretically anomalous. Muge Gocek's *Rise of the Bourgeoisie, Demise of Empire: Ottoman Westernization and Social Change* (1996) argues

---

[2] Kiser and Hechter seem to have struck something of a raw nerve among comparative historical sociologists. Quadagno and Knapp (1992, p. 500) accuse them of "epistemological terrorism," and Somers (1998) accuses them of both "ahistoricism" and "incoherence." Both the Quadagno and Knapp as well as the Somers articles are essentially extended responses to Kiser and Hechter Kiser and Hechter do set up something of a straw man in their analysis by accusing historical sociologists of a "historicism," which ignores questions of causation when, as both Quadagno and Knapp and Somers point out, the real issue is whether or not notions of causation depend on general theoretical perspectives On this point, there is a real difference of opinion.

that historical change in the semiperiphery reflects not the universal laws of a Wallersteinian world-system but rather the particular but nonetheless generalizable patterns of Ottoman history. I argue that the comparative historical practice of these second-generation scholars shows that historical particularism can play a key role in constructing social theory of a particular kind—what will be called here historically conditional theory.

## HISTORICALLY CONDITIONAL AND UNIVERSAL THEORY

Much of the debate on comparative historical method rests on the exact meaning of "general theory." Although Burawoy as well as Kiser and Hechter are united in their criticism of historical particularism, they actually have very different views of the role of theory in historical research. For Kiser and Hechter, "general theory" is a synonym for universal theory—theoretical generalizations valid in all times and places. For example, they cite the following proposition derived from rational choice theory: "the degree to which rulers are dependent on subjects for revenue and other resources" affects the relative power of rulers (Kiser and Hechter 1991, pp. 19–20). No further specification is made as to historical or cultural context, and their discussion makes clear that this proposition is made to apply to the rulers of all states from ancient empires to advanced industrial societies. Burawoy, on the other hand, cites the following proposition from Trotsky's theory of permanent revolution: "Without the direct State support of the European proletariat the working class of Russia cannot remain in power and convert its temporary domination into a lasting socialist dictatorship" (quoted in Burawoy 1989, pp. 787–88). Here both time (early 20th century) and place (Russia and Europe) are both inherent in the theoretical statement, even though the goal is to generalize beyond the Russian experience to world revolution.

Kiser and Hechter are reaffirming sociology's traditional search for "universal" laws that apply in all times and places. Burawoy, no less committed to using history to test falsifiable propositions derived from theories of society, nevertheless has an idea of theory much more closely tied to time and place. Trotsky's proposition is historically (and culturally) conditional. It does not apply to all revolutions, much less all states. On the other hand, Trotsky clearly believed that his theories of revolution applied outside Russia. His theory of "combined and uneven development" specified a set of historical conditions, similar to those in Russia during 1905–17, in which revolution was particularly likely. Trotsky's propositions are historically conditional. Conditional propositions may be derived from theoretical frameworks like Trotsky's Marxism but require specification of the historical and cultural conditions under which the proposition is believed to hold.

The term "historically conditional theory" simply makes explicit what is often implicit in theory construction. Stinchcombe's classic work on theory construction, for example, defines a causal law as "a statement or proposition in a theory which says that there exist environments (*the better described the environment the more complete the law*) in which a change in one variable is associated with a change in another variable" (Stinchcombe 1968, p. 31; emphasis added). Environment specification should, but often does not, include the historical conditions in which the law holds. Specifying these conditions involves more than simply stating the proper names of the places and periods in which the law applies. Rather, it requires, as Stinchcombe argued elsewhere (1978, p. 22), defining the "deep causal analogies" that unite situations in which the law applies.

Historically conditional causal laws may be derived from theoretical perspectives or from frameworks like Marxism but only if such perspectives are not regarded as universal. For example, in Marxism, class can be regarded as a historically specific social (and analytical) category dating from the late 19th century and useful in understanding workers' movements, or it can be understood as a universal principle of world history (the history of all hitherto existing society is the history of class struggles), in which case it is a universalizing theory that does not permit historically specific causal propositions.[3] Kiser and Hechter use the term "general theory" for what we are calling "universal theory." To avoid confusion, the term "universal" will be used here to describe theories without time space specification, and historically conditional theory will be used for theories that explicitly specify the range of historical conditions under which the theory is thought to apply. Historically conditional theory, however, unlike conjunctural or contingent explanation, does involve generalization beyond the individual case, if not to all cases, and contributes to the development of conditional theoretical frameworks, not simply to the understanding of particular historic conjunctures.

Sewell (1996) has persuasively argued that universal theories characterized the work of such first-generation scholars as Tilly, Wallerstein, and Skocpol. It will be argued that the second-generation authors reviewed here are developing neither case-specific conjunctural explanation nor universal theory but rather historically conditional theory. They do so by examining anomalies in theoretical frameworks presented by particular time-place conjunctures. In so doing, they are resolving in practice the artificial antinomy between historical specificity and theoretical general-

---

[3] For a contemporary attempt to construct a historically conditional version of Marxism, see Erik Olin Wright, Andrew Levine, and Elliot Sober, *Reconstructing Marxism* (esp. chap. 5, pp. 89–100, "Towards a Reconstructed Historical Materialism").

ization that underlies current epistemological debates over narrative, comparison, and general theory.

## *REDS OR RACKETS?* NARRATIVE, CONJUNCTURE, AND CONDITIONAL GENERALIZATION IN HISTORICAL SOCIOLOGY

Howard Kimeldorf's study of the rise of the radical ("red") West Coast International Longshoremen's and Warehousemen's Union (ILWU) in the first half of the 20th century has been used by William Sewell, Jr., as a principal exemplar of what he calls "eventful" sociology. "Kimeldorf's book, as I read it," says Sewell (1996, p. 270), "provides a potentially generalizable model of explanation or interpretation in historical sociology," because it recognizes the "path dependent, causally heterogeneous and contingent nature of temporality." In Sewell's view, Kimeldorf's narrative time represents a sharp break with the "abstract transhistorical time" of Tilly and Wallerstein and the inductive "experimental time" of Skocpol. Kimeldorf's book could equally well illustrate Steinmetz's "contingent, conjunctural causality," Griffin's "narrative explanation," Somers's "relational realism," or Abbott's "case narrative approach." Sewell argues that the historically contingent "eventful" causation of Kimeldorf is the only alternative to the universalizing generalizations of Tilly, Wallerstein, and Skocpol and of much conventional sociology.

Although Kimeldorf's study is based on a controlled comparison with the conservative racket-ridden East Coast longshoremen's union and contains conventional sociological variables, theoretical generalization is not his goal. According to Kimeldorf, it is the particular combination (conjuncture) of his variables and historical sequence of events (narrative) that leads to the radicalism of the ILWU. He isolates a series of variables on which the East and West Coast unions differ. The West Coast workers were recruited from the radical IWW-dominated culture of the maritime and logging industries; the East Coast workers came from the conservative rural Catholic cultures of Italy and Ireland. The West Coast unions faced an intransigent and concentrated three-firm oligopoly of ship owners; the East Coast owners were more numerous and divided. The Communist Party on the West Coast abandoned the party line and worked within the longshore union; the East Coast union followed the party and adopted dual unionism and, later, support for the Soviet Union, fatally isolating radicals from the union. Each of these variables—working-class culture, employer concentration, party dogmatism—could easily be treated as a causal variable in its own right and generalized to other settings, but this is not Kimeldorf's (or Sewell's) point. According to Kimeldorf, it is the "explosive combination of syndicalism, ship owner

hostility, working class resistance and communism" that molded the radical union and its leader, Harry Bridges.

These variables came together in the great San Francisco strike of 1934 that forged a generation around Bridges's leadership in a bloody waterfront confrontation with the intransigent shipowners and police. An event, the 1934 strike, itself becomes a causal variable. In fact, the 1934 strike and its role in solidifying Bridges's radical leadership are Sewell's (1996, p. 271) and Kimeldorf's strongest evidence for the contingent nature of historical causality and against sociological generalization. "An outstanding leader in the epic 1934 West Coast strike, and in the San Francisco general strike, Bridges rose, with Communist support, from ordinary dockworker to president of the San Francisco local and then, in 1936 to president of the entire Pacific coast district of the International Longshoremen's Association (ILA). The next year, he led 17,000 West Coast dockworkers out of the ILA into the CIO to form the International Longshoremen's and Warehousemen's Union (ILWU)."

This quotation, however, is not from Kimeldorf's book or Sewell's article but from a comparative study of 38 CIO unions by Judith Stepan-Norris and Maurice Zeitlin published in the *American Sociological Review* (1989, p. 511). It is used to illustrate one of their principal findings—that radical Communist activists were more likely to achieve leadership positions in CIO unions when they were part of a worker "insurgency from below" than when they established separate Communist unions. The odds were about seven times greater that a union would be Communist led if it were formed by such an insurgency than when it was not. They conclude that "concrete intraclass struggles can create . . . new objective political relations" (Stepan-Norris and Zeitlin 1989, p. 518), a conclusion consistent with Kimeldorf's argument but presented as a (conditional) theoretical generalization rather than a historical conjuncture.

Kimeldorf is a student and collaborator of Zeitlin's (as is Stepan-Norris), and his book was written as a doctoral dissertation under Zeitlin's direction. Kimeldorf's book and the Stepan-Norris and Zeitlin article are part of the same general research program on the autonomous effects of political and organizational factors on working-class consciousness and mobilization. Stated in conditional form, Kimeldorf is arguing that, when a militant worker culture confronts concentrated intransigent employers in an economic crisis like the Great Depression, union radicalism is more likely if Communist leadership breaks with party dogmatism and joins the insurgency.

These conditions are not unique to one union (the ILWU) in one historical conjuncture (the San Francisco waterfront in 1934) as Stepan-Norris and Zeitlin show; nor are they unique to the United States. Further down

the Pacific Coast, but a cultural world away from the California ILWU, in Central America in the 1930s, Communist strategy played a similar role in union insurgencies. Faced with a historic opportunity provided by a militant rural proletariat pushed to desperation by the depths of the Great Depression, and opposed by a unified and intransigent group of employers in the coffee industry, the Communist Party of El Salvador dogmatically rejected coalition with the "social fascists" of the democratic trade unions, or anyone else for that matter, and launched an insurrection entirely on its own. The party's defeat, and the subsequent massacre of most of its followers, was the defining event in modern Salvadoran political history. It extinguished the left, indeed almost any cultural memory of the left, for another 50 years.[4]

In Costa Rica, Communists faced with a similar opportunity among militant banana workers and longshoremen, and confronting an intransigent and unified employer class (there was only one—the United Fruit Company), explicitly rejected Communist orthodoxy and led the workers in their own great strike of 1934. Like the one in San Francisco in the same year, the strike defined the consciousness of a generation and made the Communist Party and its leader Manuel Mora Valverde a force in the Costa Rican banana worker's union, and the union movement generally, for the next 50 years.

Kimeldorf, of course, neither makes any explicit comparison to other unions at home or abroad nor draws any sociological (as opposed to epistemological) conclusions in his final chapter. Indeed, as the book stands, Kimeldorf fails to answer the question with which the book began—Why was the ILWU an exception to American exceptionalism, a radical union in a country in which socialism has made little headway? This question is of course directly derived from the neo-Marxian theoretical framework in which Zeitlin and his students, Kimeldorf and Stepan-Norris, work, and the failure of radicalism among American workers is a major anomaly in that theory. Kimeldorf might have proposed an original and paradoxical answer to the problem of American exceptionalism—that the Communist Party itself served to inhibit worker radicalism—but in his commitment to conjunctural causation, he ignores this potentially important theoretical implication of his own work.

On the other hand, Kimeldorf's careful mastery of historical time and place describes the historical conditions in which a causal law operates in such a way as to undercut attempts to create universal generalizations valid in all times and places. Indeed, defining the set of possible instances to which a conditional theory applies is in itself a complex theoretical and

[4] For an overview of the role of the Communist Party in El Salvador and Costa Rica in the 1930s, see Paige (1997, pp. 99–152)

empirical issue. Even if Kimeldorf's model applies to other unions in the United States and Central America in the 1930s, the exact range of historical instances to which the theory might apply would still remain to be determined. But it is precisely the role of historically conditional theory to focus attention on such questions.

Like Kimeldorf's specific causal propositions, the version of the Marxian theoretical frame employed by Zeitlin, Stepan-Norris, and Kimeldorf is clearly historically conditional, not universal; they have abandoned any commitment to such notions as the inevitable and universal triumph of the proletariat or the leading role of the Communist Party. The vanguard party has a role to be sure, but it is as likely to inhibit as to lead class-based mobilization. The triumph of socialist revolution is not assured—it depends on specific historical circumstances. A historically conditional Marxian theoretical frame generates historically conditional causal propositions. Kimeldorf's work indicates not that sociology's epic quest for causal laws is illusory, as Sewell contends, but that it has been looking for the wrong kind of laws—universal rather than conditional. The same might be said in regard to the second area of recent controversy in comparative historical epistemology—the role of comparison.

## THE ROLE OF COMPARISON, OR WHAT'S WRONG WITH THEDA SKOCPOL?

Theda Skocpol's *States and Social Revolutions* was recently chosen as one of the 10 most influential books in sociology in the last 25 years and ranks as one of the top sellers among books in sociology according to *Contemporary Sociology* (May 1996, p. 293; March 1977, p. 134). Its goals are as ambitious epistemologically as they are theoretically. It is in fact a methodological manifesto of the new sociology of comparison. Skocpol contends that inductive cross-societal comparison following Mill's "method of difference" and "method of similarity" is the way in which comparative historical sociology should be practiced, and in *States,* she does just that, comparing three "similar" cases of "social revolution" (France, Russia, and China) with four other cases that "differ" in that they did not have successful social revolutions, even though they had some of the conditions that might be expected to produce them (England 1640–89; Germany 1807–14 and 1848; Japan 1868).

As Skocpol explains Mill's methods:

> Basically one tries to establish valid associations of potential causes with the given phenomenon one is trying to explain. There are two main ways to proceed. First, one can try to establish that several cases having in common the phenomenon one is trying to explain also have in common a set of causal factors, although they vary in other ways that might have seemed

causally relevant. This approach is what Mill called the "Method of Agreement." Second, one can contrast the cases in which the phenomenon to be explained and the hypothesized causes are present to other cases in which the phenomenon and the causes are both absent, but which are otherwise as similar as possible to the positive cases. This procedure Mill labeled the "Method of Difference " (Skocpol 1979, p. 36)

## Skocpol in the Critic's View: Induction and Comparison without Theory or History

Skocpol's application of Mill's method has been the subject of extensive commentary and debate (Burawoy 1989; Kiser and Hechter 1991; Lieberson 1992; Nichols 1986; Ragin 1987; Sewell 1996; Tilly 1984), and the emerging consensus is that the method itself, at least as interpreted by Skocpol, is seriously flawed. The criticism are of several types. First, there has been the issue of numbers of "cases" (Kiser and Hechter 1991; Lieberson 1992; Ragin 1987). With such a small number of cases and so many potential variables, it is argued, there is no possibility that any purely inductive approach will lead to any definitive results. Second, critics have rejected the idea of pure induction itself as fatally flawed (Burawoy 1989; Kiser and Hechter 1991; Sewell 1996) because, as Kiser and Hechter (1991, pp. 14–15) argue, "some theory is also necessary to guide the choice of appropriate cases and to select appropriate factors for inclusion in the model. The methods of agreement and difference will produce results that are a direct function of the cases chosen and the factors included."

Third, critics have argued that even if the Millian method is valid, Skocpol does not practice it very well (Burawoy 1989; Nichols 1986; Sewell 1996; Tilly 1984). According to these critics, Skocpol's cases of social revolution lack one or more of the variables her theory calls for, and her contrasting cases have some aspects of social revolution present without the theoretically necessary prerequisites. The critics also argue that Skocpol fails to disprove alternative theories, or even to consider them seriously (Tilly 1984, p. 114). According to her critics, Mill's method, if consistently applied, does not disprove Skocpol's theory and fails to disprove the principal alternative theories.

Finally, critics have argued that, despite her claims to comparative historical innovation, Skocpol has in fact ignored history at least three ways. First, by choosing "cases" of "social revolution" in such disparate time periods (spanning three centuries), she has ignored the most fundamental argument of historical sociologists—that the effects of causes differ in different epochs (Burawoy 1989; Sewell 1996). Second, by lumping together wildly different historical events under a common rubric of "social revolution," she has ignored their very real historical differences, including, as

Sewell (1985) has argued, the not inconsequential fact that their ideologies and ideological consequences are different. Finally, she has ignored historical time, sequence, and narration by making cross-sectional comparisons and not incorporating sequence in any systematic way into her theory (Burawoy 1989; Kiser and Hechter 1991; Sewell 1996).

These critiques of the most prominent advocate of the Millian method have led some of the critics to conclude that the Millian method of comparison is itself flawed. Skocpol herself (1986, 1994) has vigorously rejected most of the criticism of her and her method. My own view is that the critics are responding not so much to Millian comparison but to Skocpol's search for universal causal laws through an inductive search for common elements in a population of cases. Millian methods are entirely appropriate in historical comparative work if they are used to examine anomalies in historically conditional theoretical frames and develop conditional, not universal, causal propositions. Gay Seidman, in her book *Manufacturing Militance* (1994), in my opinion, does just that.

### Conditional Theory, Millian Methods, and Marc Bloch's "Natural Experiment": Seidman's *Manufacturing Militance*

Seidman sets out explicitly to practice Mill's "method of similarity" and explicitly acknowledges Skocpol as a model (Seidman was an undergraduate student of Skocpol). "Rather than contrasting different outcomes in cases that might have similar backgrounds [Mill's method of difference], it is also possible to use comparisons to try to explain similar outcomes in different contexts, exploring the common dynamics that shape social phenomena [Mill's method of agreement]" (Seidman 1994, p. 9). A comparison of workers' movements in Brazil and South Africa, she says, "begs the latter type of comparison" because of their "very different cultures" (p. 9). These differences include most notably, of course, the system of state-sponsored apartheid in South Africa, which Skocpol (1979, p. xii), by her own account, regarded as the reason no revolution had occurred "or likely soon would" in that country.

It is surprising, therefore, to find that the labor movement as it developed in the two societies between 1970 and 1985 took on remarkably similar forms. In both cases, industrial workers in mining and manufacturing participated in a highly politicized union movement that also mobilized many people in the informal economy in a pattern Seidman calls "social movement unionism." Demands for the democratic opening of closed authoritarian systems were an important part of both movements. Remarkably, given the intersection of race and class in both societies and the institutionalized racism of South Africa, both movements were notably

nonracist (or "racialist" as they would say in South Africa). Class and politics transcended purely national or racial appeals.

Whatever it was that caused social movement unionism, it obviously could not have been apartheid since that was present in only one of the two cases. Nor could it have been either Anglo-Dutch or Luso-African cultural traditions, since these two variables differed in the two cases. It must have been something that the two societies shared. For Seidman, the answer was that both were semiperipheral societies undergoing rapid industrialization and dependent development under the direction of authoritarian states during a period of worldwide economic growth, followed by a sharp economic slowdown (after 1973). The success of the bureaucratic authoritarian industrializing strategies of both Brazil and South Africa created a large and poorly paid industrial proletariat (low wages were a strategy of both regimes to enhance their competitive positions in the world economy). The narrow economic focus of both regimes relegated much of the population to the urban squatter settlements that are such characteristic features of Brazilian and South African cities. To make matters worse, both governments neglected even minimal welfare services for these settlements.

Both regimes were infamous for their overdeveloped security apparatuses, death squads, and efficient elimination of their political opponents. Under these circumstances, social movement unionism, or indeed any opposition, might seem doomed to failure. But the economic downturn after 1973 provided an opportunity, while at the same time dramatically increasing the economic pressure on both workers and the urban poor. The business elite, united behind the repressive apparatus in the past, began to see its interests diverging as growth slowed and competition with the state for scarce resources increased. The security apparatus, which had seemed a necessity in the past, now seemed an obstacle to further growth and stability. Militant social movement unionism then split the elite with the business sector moving toward accommodation and withdrawing its support from authoritarian rule leading to nearly simultaneous transitions to democracy in both countries.

In constructing this argument, Seidman managed to avoid the alleged pitfalls of the Millian method emphasized by the critics. First, she ignored Skocpol's strictures about "the blinders, or the heavily tinted lenses, of pre-existing theories" (1979, p. 9) and explicitly set out to explain an important theoretical anomaly. It is in fact exactly the same anomaly in Marxian theory that puzzled Kimeldorf—the failure of revolution among the industrial proletariat. Like his dockworkers, the workers of Brazil and South Africa are the exception to the exception—militant politicized industrial workers. Although they stopped far short of revolution, they contributed to the demise of authoritarian governments in both cases. Seid-

man also notes, however, that her cases are an anomaly from the point of view of modernization theory since they fail to follow the route of political incorporation of Western European workers as described by Bendix (Seidman 1994, p. 258). Both her cases and variables were theoretically selected.

Second, she does not ignore historical conditions but makes them an important part of her argument. She focuses her attention on the same historical time period—the peak of industrial expansion in the postwar era and the downturn after 1973—and she follows two societies that have remarkably similar economic trajectories. The historical phenomena under study really are similar enough to suggest that they can be treated as equivalent for analytical purposes. Social revolutions in France in 1789 and Russia in 1917 do not look much alike to most observers, but much of the power of Seidman's analysis depends on just how unexpectedly similar her two union movements are. The industrial policies of the two governments are also remarkably similar. A temporal sequence is followed in both cases. Indeed, the sequence itself, especially the transition to democracy, is an important part of her theory.

Seidman is also demonstrating that good comparative work does not depend in any linear way on the number of so-called "cases" examined. Skocpol's evocation of inductive empiricism has done a disservice to comparative historical sociology by framing the debate along the "large $N$"/ "small $N$" division familiar in positivist social science. In fact, as Griffin (1992) has argued, Skocpol shares the same inductive logic as statistical sociology, differing only in the size of the $N$. Fortunately, Seidman ignored Skocpol's ideas on inductive case logic and simply went ahead and tested theoretical ideas in her two "cases." Perhaps we need to look once again at the methods of comparative inquiry as described by the great historian Marc Bloch (1953) who argued that a carefully framed hypothesis developed in one context could be tested by examining it in another context in which similar conditions might be expected to prevail. Even a single observation ("case") could serve as a test in what for Bloch represented a "natural experiment."

As Ragin (1987, pp. 61–64) points out, the logic of such natural experiments differs fundamentally from statistical analyses of large numbers of cases. Statistical analysis requires that the analyst assume that ($a$) causes work the same way in each case, independent of the presence of other causes in that case; ($b$) effects are linear and additive; and ($c$) the average effect of each cause in a population can be applied to every case in the population. The net effect of these assumptions is to produce estimates of association in populations whose theoretical implications, as Ragin (1987, p. 63) notes, are "of unknown value." As Lieberson (1985, p. 103) has pointed out, there is no necessary relation between such measures of

strength of association and causal statements.[5] Turner (1987) and Freedman (1991), as well as Meehl (1970), point to the same difficulties. Although these defects can be remedied by running analyses on subgroups, by introducing "interaction terms," or by analyzing residuals, these techniques can reduce, but not eliminate, the averaging properties of causes across cases inherent in statistical methods.

So, in general, statistical analysis proceeds by ignoring anomaly and describing average probabilities in populations as universal "causal laws," even though these probabilities may in fact apply to *no case* in the population. And attempts to remedy these defects through interaction terms or analyses of subpopulations will rapidly exhaust the multiple cases that are the basis for statistical analysis. As the critics point out, Skocpol's adoption of a similar inductive approach lacks a sufficient number of cases even to assess average effects. Bloch's natural experiment, like Burawoy's anomaly, on the other hand, focuses on theoretically defined historical conjunctures, which may easily be a *single* observation or "case."

At the most fundamental level, the contrast between Seidman's and Skocpol's use of the Millian method is between an inductive quasi-statistical search for universal laws that apply to all social revolutions and a consideration of anomalies in historically specified cases. Seidman's generalizations, unlike Skocpol's, are historically conditional. The conclusions apply to the strategies pursued by late developers during the economic crisis beginning in the 1970s. In her conclusions, Seidman tentatively and cautiously applies her model to the Korean case and suggests that there may be parallels in other recent industrializing states but rejects universal generalizations to all "late industrializers," regardless of historic period.

This is not to say that Seidman is not working within and contributing to a more general theoretical framework. Seidman, like Kimeldorf, is testing a historically conditional version of Marxian theory in which "different capitalist growth patterns will produce different experiences, different possibilities for labor, different working class movements" (1994, p. 274), not the inevitable triumph of the proletariat. In both cases, their findings have implications for a historically conditional Marxian theoretical frame. In neither case is Marxism treated as a universal theory.

---

[5] Lieberson (1985, pp. 99–107) points out that a hypothetical statistical study of the law of gravity based on observations of falling objects (coins, feathers, etc.) through air would find that their rate of descent depended on their density, surface area, shape, and so on, but would ignore gravity because it is a constant that does not vary in the population of observed values.

## UNIVERSAL AND CONDITIONAL THEORY: OUT OF THE WALLERSTEINIAN ORRERY

As Kiser and Hechter (1991, p. 3) argue, much of the resistance to generalizing social theory in comparative historical work is in fact a reaction to the perceived failures not so much of theory but of the particular universal theories that have organized sociological thought since its 19th-century origins—functionalism from Durkheim and Malinowksi to Parsons and Merton; Marxism from Marx himself to the neo-Marxian theory of Wallerstein; and versions of modernization theory from Tönnies to Inkeles. There is now a general distrust of all such "metanarratives" that purport to discover general laws of history, probably for good reason. Underlying all of these grand theories of history were universal, teleological notions of progress. In such theories, all societies were governed by the same universal laws and followed the same sequence toward the ultimate (or, in Marxism, penultimate) stage of modern industrial society.

One alternative to these now discredited universal theories is of course a conjunctural causality, which denies that the construction of causal law or general theoretical frameworks is possible. But as we have seen, throwing out universality and teleology does not necessarily imply abandoning causal propositions or theoretical frameworks. On the contrary, it is a necessary first step in the construction of historically conditional theory. This distinction is well illustrated by the contrast between the explanations of a particular historical case, the fall of the Ottoman empire, according to one of the most influential of all universal comparative historical theories—Wallersteinian world-system theory—and according to the historically conditional theory of Muge Gocek in her *Rise of the Bourgeoisie, Demise of Empire.*

The Ottomans present no real problems for Wallerstein's world-system theory since, after some minor resistance, they fell into their proper place in the expanding world economy in the period 1730–1840, becoming a peripheral exporter of primary products after European manufactures flooded their markets and strong European states pressured the empire to eventual collapse (Wallerstein and Kasaba 1982, cited in Gocek 1996, p. 17; Wallerstein 1979, 1989). The Wallersteinian orrery whirls around its European center undisturbed by the perturbations of Ottoman history. As Howe and Sica (1980, p. 225) observe, the system's "future is inscribed in its conception." Since there is a core, there must be a periphery, and the Ottomans must be in it. Why the Ottomans's response to the West, like Japan's, could not just as well have led them into the core is not explained in world-system theory. The question of which societies end up in the core and which end up in the periphery is of little theoretical interest.

The question of who is in the core and who is in the periphery is in fact the starting point for Muge Gocek's analysis. What is unproblematic for Wallerstein is an anomaly for her—why did the incipient Ottoman bourgeoisie not follow the same path as its Western rivals from agrarian to industrial society? Turkey was a puzzling anomaly for Marx as well, contributing to his formulation of the notorious "Asiatic" mode of production as a conjunctural if "Orientalist" explanation. Wallerstein's response to the same anomaly was to stand Marx's teleology on its head and declare that peripheralization, not industrialization, was the expected outcome. For Gocek, neither the transition from feudalism to industrial capitalism nor from the "Asiatic" mode to the periphery is inevitable, and the failure of bourgeoisie modernization is, therefore, an anomaly.

Gocek's attempt to explain this anomaly begins where Wallerstein leaves off—with the impact of European capitalism on the Ottoman Empire. But from there, the analyses diverge markedly. According to Gocek, the Ottoman sultans made heroic efforts to meet the Western challenge, particularly by modernizing their military forces through the import of Western military technology, Western advisors, and Western schools. Although the effort was only a modest success in its intended aims, it did succeed in creating what Gocek calls a bureaucratic bourgeoisie—a Western-educated, secular elite of military officers and bureaucrats whose influence has been felt in Turkish society ever since. In the end, they became the sultan's most effective opponents and ultimate grave diggers.

Western commercial influences transformed Ottoman society just as Wallerstein said they would. But because of the peculiarities of Ottoman imperial policy, Islamic citizens were in a poor position to take advantage of these developments. The sultan regarded independent centers of commercial wealth as a challenge to his rule and increasingly relegated them to the control of powerless minorities—Christian Armenians and Jews who were well suited to take advantage of commercial contacts with their coreligionists in Western Europe. As "Westernization" advanced and the crisis of the Ottoman Empire deepened, the emerging bourgeoisie segmented into a Westernized but ethnically Turkish bureaucratic segment in charge of the army and a commercially powerful but politically weak, ethnically distinct, economic segment.

When military pressure finally brought the empire to its knees in World War I, the two halves of the elite split apart with disastrous consequences for the both the sultan and the emerging Turkish nation-state. The modern Turkish nation was founded by the modernizing, secular, bureaucratic bourgeoisie on the ashes of the Ottoman Empire. The fate of the economic bourgeoisie was liquidation, forced migration, and, in the case of the Armenians, physical extinction. At one and the same moment, the Turks founded a nation-state and destroyed the capitalist bourgeoisie that might

have made Turkey a modern industrial power. So Turkey became then what it remains today, a part of Wallerstein's "semiperiphery" with a strong state and a weak economy.

Neither the rise of Turkish capitalism nor its demise is seen as some master teleological plan as it was in Marx's transition from feudalism to capitalism or Wallerstein's peripheralization. But this does not mean that there is only conjunctural explanation in Gocek's work. Like Seidman (but unlike Kimeldorf), Gocek attempts to generalize her conclusions to other cases that share the same set of conditions as the Ottoman empire— a historically created division between a bureaucratic bourgeoisie and an ethnically distinct commercial bourgeoisie. She mentions explicitly Kenya, Malaysia, and Uzbekistan (1996, p. 141) as societies that may share these conditions and suggests that they may be widespread in Africa and the Middle East. Her propositions are not, however, universal. They do not necessarily apply to the historic experience of the West or of Latin America, for example, nor even to all societies in the Middle East and Africa.

Gocek's work also indicates that historical investigations based on a consideration of theoretical anomalies or Bloch's natural experiment do not necessarily require more than a single "case." Even though Gocek, unlike Kimeldorf or Seidman, is studying only one "case," this does not imply that her study is "only a case study." As John Walton (1992, pp. 121–23) has argued, the term "case study" itself is a rhetorical claim that the "case" is an exemplar of a "universal" process that can only be revealed by statistical analysis. Gocek, Seidman, and Kimeldorf, on the contrary, argue that their "case(s)" are not exemplars but anomalies. This difference in the treatment of the "case" is one of the clearest indications of the fundamental epistemological distinction between statistical and comparative historical analysis. For statistical analysis, the goal is to find typical or average effects in populations so a case can at best be an illustration of such effects. For comparative historical analysis, the goal is to find causal principles, and a case is regarded, after Mill or Bloch, as a natural experiment in which competing causal theories can be tested.

World-system theory shares with statistical analysis this concern with universal social process and neglect of anomalies. Wallerstein's attempt to demonstrate, for example, that the *same* process of incorporation characterized the Ottomans, the Russian Empire, India, and West Africa leads him to equate their mid–19-century states as "neither too strong, nor too weak" (1989, p. 187), even though India and West Africa were formal colonies, Russia an imperial power in its own right, and the Anatolian core of the Ottoman empire was never colonized. Similarly, Wallerstein's inability to account for variation in response to capitalist penetration is, as Gocek (1996, 152 n. 71) and a number of other critics have pointed out

(Brenner 1977; Chirot and Hall 1982; McMichael 1990; Stinchcombe 1982), a crucial weakness of his theory. As Skocpol (1977, pp. 1088–89) noted in her review of the first volume of *The Modern World-System*, in Wallerstein's universal style of argument, "deviant historical cases do not force one to modify or replace one's theory."

CONCLUSIONS

If, as Burawoy argues, the recognition and resolution of anomalies is the way in which a research program progresses, then *neither* statistical generalization nor world-system or other universal theory has this capacity for theoretical advance. The successful resolution of an anomaly will have the structure of what Stinchcombe (1968, p. 24) calls a "crucial experiment"—one that disconfirms one theory while increasing confidence in another. A strong aesthetic reaction to a crucial experiment, is, according to Stinchcombe (p. 28), the "mark of the true scientist." The work of Kimeldorf, Seidman, and Gocek generates such a reaction because it simultaneously challenges received theoretical wisdom while providing convincing evidence for an alternative theoretical conception. This is exactly what Marc Bloch said the comparative method was supposed to accomplish.

There would of course be no anomalies without theoretical frameworks so that conjunctural explanation does not have this property to generate crucial experiments and hence advance social theory. Even theories that appear to be universal and teleological may generate theoretical propositions if the processes they describe are treated as historical rather than inevitable. Gocek, Kimeldorf, and Seidman are all working through anomalies in a particular theoretical framework—Marxism (including Wallerstein's neo-Marxism) often regarded as universalizing. In all three cases, however, theoretical propositions derived from Marxism are regarded as historically contingent, not universalizing or teleological.

While historical sociologists have been debating the status of theory in comparative historical research, researchers have been ignoring the debates and effectively using theory to do historical work that combines narrative and conjunctural temporality, theory-driven comparison, and historically conditional but generalizable propositions. At the same time, they have demonstrated that it is possible to contribute to the construction of a historically conditional version of a theoretical framework, without losing sight of the particulars of time and place. The work of these second-generation scholars indicates there is no contradiction between comparative historical work and sociological theory. On the contrary, comparative historical work is essential for the development of conditional as opposed to universal sociological theory.

## REFERENCES

Abbott, Andrew. 1990. "Conceptions of Time and Events in Social Science Methods: Causal and Narrative Approaches." *Historical Methods* 23 (Fall): 140–50.

———. 1992*a*. "From Causes to Events: Notes on Narrative Positivism." *Sociological Methods and Research* 20:428–55.

———. 1992*b*. "What Do Cases Do? Some Notes on Activity in Sociological Analysis." Pp. 53–82 in *What Is a Case?* edited by Charles Ragin Cambridge: Cambridge University Press.

Aminzade, Ron. 1992. "Historical Sociology and Time." *Sociological Methods and Research* 20:456–80.

Bloch, Marc. 1953. "Toward a Comparative History of European Societies." Pp. 494–521 in *Enterprise and Secular Change,* edited by Frederic C. Lane and Jelle C. Riemersma Homewood, Ill.· Irwin.

Brenner, Robert. 1977. "The Origins of Capitalist Development: A Critique of Neo-Smithian Marxism." *New Left Review* 104:25–92.

Burawoy, Michael. 1989. "Two Methods in Search of Science: Skocpol versus Trotsky " *Theory and Society* 18:759–805.

Chirot, Daniel, and T. D. Hall. 1982. "World-System Theory." *Annual Review of Sociology* 8:81–106.

Freedman, David 1991. "Statistical Models and Shoe Leather." *Sociological Methodology* 21·291–313

Gocek, Muge. 1996. *Rise of the Bourgeoisie, Demise of Empire: Ottoman Westernization and Social Change.* New York: Oxford University Press.

Griffin, Larry 1992. "Temporality, Events and Explanation in Historical Sociology." *Sociological Methods and Research* 20:403–27.

———. 1993. "Narrative, Event-Structure Analysis, and Causal Interpretation in Historical Sociology." *American Journal of Sociology* 98:1094–133.

Howe, Gary N., and Alan M. Sica. 1980. "Political Economy, Imperialism, and the Problem of World-System Theory." Pp. 235–86 in *Current Perspectives in Social Theory,* edited by Scott McNall. Greenwich, Conn.: JAI Press.

Kimeldorf, Howard 1988. *Reds or Rackets? The Making of Radical and Conservative Unions on the Waterfront.* Berkeley and Los Angeles: University of California Press.

Kiser, Edgar, and Michael Hechter 1991 "General Theory's Role in Comparative-Historical Sociology." *American Journal of Sociology* 97:1–30.

Lieberson, Stanley, 1985. *Making It Count: The Improvement of Social Research and Theory.* Berkeley and Los Angeles: University of California Press.

———. 1992. "Small *N*'s and Big Conclusions An Examination of the Reasoning in Comparative Studies Based on Small Numbers of Cases." Pp. 105–18 in *What Is a Case? Exploring the Foundations of Social Inquiry,* edited by Charles C. Ragin and Howard S. Becker. Cambridge. Cambridge University Press.

McMichael, Philip. 1990. "Incorporating Comparison within a World-System Perspective: An Alternative Comparative Method." *American Sociological Review* 55:385–97.

Meehl, Paul A. 1970 "Nuisance Variables and the Ex Post Facto Design." *Minnesota Studies in the Philosophy of Science* 4·373–402.

Nichols, Elizabeth. 1986 "Skocpol on Revolution Comparative Analysis vs. Historical Conjuncture." *Comparative Social Research* 9:163–86

Paige, Jeffery M. 1997. *Coffee and Power: Revolution and the Rise of Democracy in Central America.* Cambridge, Mass : Harvard University Press.

Quadagno, Jill, and Stan J. Knapp. 1992. "Have Historical Sociologists Forsaken Theory? Thoughts on the History Theory Relationship." *Sociological Methods and Research* 20:481–507.

Ragin, Charles. 1987. *The Comparative Method: Moving Beyond Qualitative and*

*Quantitative Strategies.* Berkeley and Los Angeles: University of California Press.

Seidman, Gay. 1994. *Manufacturing Militance: Workers' Movements in Brazil and South Africa 1970–1985.* Berkeley and Los Angeles: University of California Press.

Sewell, William H. 1985. "Ideologies and Social Revolutions: Reflections on the French Case." *Journal of Modern History* 57 (3): 81–84.

———. 1996. "Three Temporalities· Toward a Sociology of the Event." Pp. 245–80 in *The Historic Turn in the Human Sciences,* edited by Terence J. McDonald. Ann Arbor: University of Michigan Press.

Skocpol, Theda. 1977. "Wallerstein's World Capitalist System· A Theoretical and Historical Critique." *American Journal of Sociology* 82:1075–90.

———. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia and China.* Cambridge: Cambridge University Press.

———. 1986. "Analyzing Causal Configurations in History. A Rejoinder to Nichols." *Comparative Social Research,* 9:187–94.

——— 1994. *Social Revolutions in the Modern World.* Cambridge: Cambridge University Press.

Somers, Margaret R. 1996. "Where Is Sociology after the Historic Turn? Knowledge Cultures, Narrativity, and Historical Epistemologies." Pp. 53–90 in *The Historic Turn in the Humans Sciences,* edited by Terence J McDonald. Ann Arbor: University of Michigan Press.

———. 1998. "'We're No Angels': Realism, Rational Choice, and Relationality in Social Science." *American Journal of Sociology* 104 (3): 722–84.

Steinmetz, George. 1998. "Critical Realism and Historical Sociology " *Comparative Studies in Society and History* 40 (1) 170–86.

Stepan-Norris, Judith, and Maurice Zeitlin. 1989. " 'Who Gets the Bird?' or How the Communists Won Power and Trust in American Unions." *American Sociological Review* 54:503–23.

Stinchcombe, Arthur L. 1968. *Constructing Social Theories.* New York: Harcourt.

———. 1978. *Theoretical Methods in Social History.* New York: Academic Press.

———. 1982. "Review Essay: The Growth of the World System." *American Journal of Sociology* 87:1389–95.

Tilly, Charles. 1984. *Big Structures, Large Processes, Huge Comparisons.* New York: Russell Sage.

———. 1994. "The Time of States." *Social Research* 61:269–95.

Tomich, Dale. 1995. "Small Islands and Huge Comparisons: Caribbean Islands, Historical Unevenness, and Capitalist Modernity." *Social Science History* 18 (3): 339–57.

Turner, Stephen P. 1987. "Cause, Law and Probability." *Sociological Theory* 5.15–19.

Wallerstein, Immanuel. 1979. "The Ottoman Empire and the Capitalist World-Economy: Some Questions for Research." *Review Journal of the Fernand Braudel Center* 2:389–98.

———. 1989. *The Modern World-System, III: The Second Era of Great Expansion of the Capitalist World Economy, 1730–1840s.* New York: Academic Press.

Walton, John. 1992. "Making the Theoretical Case " Pp. 121–37 in *What Is a Case? Exploring the Foundations of Social Inquiry,* edited by Charles S. Ragin and Howard S. Becker Cambridge: Cambridge University Press.

Wright, Erik Olin, Andrew Levine, and Elliott Sober. 1992. *Reconstructing Marxism: Essays on Explanation and the Theory of History.* London: Verso.

# Institutional Logics and the Historical Contingency of Power in Organizations: Executive Succession in the Higher Education Publishing Industry, 1958–1990[1]

Patricia H. Thornton
*Duke University*

William Ocasio
*Northwestern University*

This article examines the historical contingency of executive power and succession in the higher education publishing industry. We combine interview data with historical analysis to identify how institutional logics changed from an editorial to a market focus. Event history models are used to test for differences in the effects of these two institutional logics on the positional, relational, and economic determinants of executive succession. The quantitative findings indicate that a shift in logics led to different determinants of executive succession. Under an editorial logic, executive attention is directed to author-editor relationships and internal growth, and executive succession is determined by organization size and structure. Under a market logic, executive attention is directed to issues of resource competition and acquisition growth, and executive succession is determined by the product market and the market for corporate control.

Classic and contemporary theory and research in sociology have emphasized the importance of leadership succession in demarcating changes in

power and authority in organizations and in society (Fligstein 1987, 1990). Among the early studies in the sociology of organizations, Gouldner (1954) and Grusky (1960, 1963) defined succession, the replacement of key officials, as a critical process in the transition from personal to bureaucratic patterns of control in organizations. Contemporary research on succession has continued to focus on the sources of executive power and authority, with an emphasis on the internal political dynamics in organizations. For example, recent studies examine the social psychology of CEO relationships, the demographic and structural characteristics of boards of directors, and the political dynamics among board members and top management (Boeker 1992; Ocasio 1994; Zajac and Westphal 1996, respectively).

Institutional approaches suggest a different focus for studies of leadership power in organizations—that interests, power, and politics in organizations are shaped by institutional logics prevailing in wider environments (Fligstein 1990; Friedland and Alford 1991; Powell 1991; Davis and Greve 1997; Meyer et al. 1997). According to this view, while power and politics are present in all organizations, the sources of power, its meaning, and its consequences are contingent on higher-order institutional logics. Institutional logics define the rules of the game by which executive power is gained, maintained, and lost in organizations (Jackall 1988). Moreover, institutional logics are historically variant and are shaped by economic and social structural changes (Fligstein 1985, 1987; Fligstein and Brantley 1992; Barley and Kunda 1992). However, the effects of institutional logics on the determination of power in organizations is not emphasized in most empirical analyses of intraorganizational power or, in particular, in recent studies of succession. While a general theme of both classic and contemporary studies on leadership succession is that organizational politics shape executive change, the idea that the political determinants of succession are themselves conditioned by historical context and institutional logics has been relatively unexplored, with the exception of Fligstein (1987).[2]

In this article, we explore how a historical shift in the dominant institutional logic in an industry from the logic of professions to the logic of markets led to a transformation in the political dynamics in organizations and the determinants of executive succession. We select the higher educa-

---

[2] Note, however, that Fligstein (1987) does not study executive succession directly but instead studies the functional backgrounds of executives at different time periods. Furthermore, he focuses on how executive power is shaped by changes in state policies (Fligstein 1990) and changes in organizational strategies and structures (Fligstein 1987, 1990) Our study differs from Fligstein's in our theoretical emphasis on the consequences of change in industry-level logics and in our empirical focus on the shift from a professional to a market logic. Note that Fligstein's (1990) conceptions of control may be considered an alternative logic to either professions or markets—with production, marketing, and financial conceptions as three distinct variants of managerialism.

tion publishing industry for the analysis because it provides a vivid case of historical variation in the prevailing institutional logics. Previous studies on publishing suggest that it experienced a transformation from an industry that emphasized publishing as a profession to one where executives paid increasing attention to the influences of competition and market forces (Powell 1985; Tebbel 1987). Combining interview data with historical analysis, we examine the characteristics of the institutional logics at different time periods, which reveal the shift from an editorial logic to a market logic. We then estimate hazard rate models of executive succession events in the industry from 1958 to 1990, analyzing how the historical changes from an editorial to a market logic shaped the relative importance of the positional, relational, and economic determinants of executive power and succession.

This article makes several contributions. First, it establishes empirically the historical contingency of the potential sources of power and executive succession in organizations. Linking neoinstitutional theory and research to the study of intraorganizational power and executive succession, we develop and test hypotheses that focus on the historical periodization of sources of executive power and control (Fligstein 1987, 1990). We extend existing studies of organizational politics by exploring how the salience of various sources of power, both internal and external to the organization—formal position and rank in the hierarchy, organization size and differentiation, ownership form, resource competition, and the market for corporate control—are historically contingent on the prevailing institutional logics. In particular, we show for the higher education publishing industry a historical increase in the importance of economic determinants of executive power and a decline in the effects of positional and relational sources of power on executive succession.

Second, this article makes a theoretical contribution by developing a set of mechanisms, operating across multiple levels of analysis, by which institutional logics shape power in organizations. At the macrolevel, we build upon Friedland and Alford's (1991) focus on institutional logics as supraorganizational patterns, both symbolic and material, that order reality and provide meaning to actions and structure conflicts. At the industry level, we propose that logics are embodied in the common identity of industry players, which is based on social comparison and status competition among competitors (Porac et al. 1995; White 1992). At the level of organizational actions and decisions, we focus on the social organization of attention and decision making (Simon [1947] 1997; Ocasio 1997; Zerubavel 1997). We extend theory by proposing three mechanisms by which institutional logics direct attention to alternative sources of power in organizations.

Third, this article extends the empirical analysis of institutional perspectives by combining interviews and historical research with the use of

piecewise exponential models to test the consequences of differing institutional logics on succession. The piecewise specification tests for interaction effects of the organizational and market determinants of power with the institutional logics prevailing in different historical periods. To our knowledge, this is the first study to combine an extensive data set using interviews, literature reviews, and quantitative modeling with annual event history data on individual actors, organizations, and their environments.

This article is organized as follows. First, we develop our theory of how a transformation in institutional logics changes the determinants of executive succession. Second, we use interviews and historical research to develop a typology of two ideal types of institutional logics: editorial and market. Third, we formulate hypotheses that relate the attributes of these ideal types to the dependent variable, executive succession. Last, we use event history models to test our hypotheses on how changes in the institutional logics affect the rate of executive succession.

## THEORY

According to political theories (Perrow 1986; Pfeffer and Salancik 1978; Pfeffer 1981, 1992), executives derive their power and authority from their formal position in the organization, their social relationships, the organization's reputation and status, and the executives' ability to manage the organization's strategic contingencies and resource dependencies. This article builds on existing power and politics perspectives by examining how these determinants of power in organizations are historically contingent. Our theory and analysis, while consistent with resource dependence theory and structural contingency perspectives, emphasizes how the determinants of power, including dependencies and contingencies, are moderated by the prevailing institutional logic (Jackall 1988; Friedland and Alford 1991; Haveman and Rao 1997, p. 1607).

We define institutional logics as the socially constructed, historical pattern of material practices, assumptions, values, beliefs, and rules by which individuals produce and reproduce their material subsistence, organize time and space, and provide meaning to their social reality (Jackall 1988, p. 112; Friedland and Alford 1991, p. 243). Institutional logics are both material and symbolic—they provide the formal and informal rules of action, interaction, and interpretation that guide and constrain decision makers in accomplishing the organization's tasks and in obtaining social status, credits, penalties, and rewards in the process (Ocasio 1997). These rules constitute a set of assumptions and values, usually implicit, about how to interpret organizational reality, what constitutes appropriate behavior, and how to succeed (Jackall 1988; March and Olsen 1989).

Institutional perspectives emphasize how historical change is important

in understanding the patterns of power and control in organizations (Brint and Karabel 1991; Fligstein 1987, 1990). This notion dates back to Weber ([1922] 1978) and his identification of historically situated ideal types: control by individual charisma, by tradition, and by legal bureaucracy. Similarly, Grusky recounts how the problem of succession went through several historical stages that accompanied institutional change in the governance of American business: (1) control by the charismatic leader—the entrepreneur-founder, (2) transfer of control on the basis of kinship, and (3) change in control on the bureaucratic basis of professional competence. He noted that "the particular sequence of stages should influence the nature of the succession problem" (Grusky 1960, p. 110). More recently, Fligstein (1987) and Ocasio and Kim (1999) found that the functional backgrounds of those who rose to power in large U.S. corporations was determined by historical shifts in conceptions of control. Brint and Karabel (1991) showed how the ability of administrators to promote change and to advance their own managerial interests in U.S. community colleges was shaped by new institutions for student testing and placement and the adoption of the guiding ideology of vocationalization.

We follow Friedland and Alford's (1991) account of how the institutions that shape organizational action are embedded within higher-order societal logics. The major institutions of society—the market, the state, the corporation, the professions, religion, and the family—provide a distinct set of often contradictory logics that form the bases of political conflicts. Friedland and Alford (1991, p. 242) argue that individuals, organizations, and society constitute three nested levels, where organization and societal-level institutions specify progressively higher levels of opportunity and constraint for individual action.

We examine an additional level at which institutional logics operate—the industry. Note that our focus on industry logics complements past research that focuses on logics embodied in organizational forms (Haveman and Rao 1997). We suggest that an industry is a relevant boundary for identifying institutional logics because industry producers develop common identities and "valuation orders" that structure the decision making and the practices of the players in a product market (White 1981, 1992). Such common identities and valuation orders emerge from social comparisons among firms under conditions of resource and status competition. These common identities and valuation orders become institutionalized particularly where such practices map to legitimating accounts (Strang and Meyer 1994; Davis and Greve 1997). Therefore, institutional logics provide an articulation between the social and economic structures and the rules and meanings that constitute a commonly understood set of actions within the industry (Porac et al. 1995).

More specifically, we explore how industry-level institutional logics

shape the positional, relational, and economic determinants of executive power and succession across different historical periods. By positional, we mean those determinants of power that are inherent in the actor's role or position, such as founder or division executive. By relational, we mean those determinants of power that derive from the structure of relationships among actors, and groups of actors both intra- and interorganizationally. By economic, we mean the determinants of power related to issues that have consequences for firms in a product market and their performance, such as access to public capital markets, resource competition, profitability, and the market for corporate control.

We propose three mechanisms by which industry-level institutional logics shape executive power in organizations. First, the meaning, appropriateness, and legitimacy of various sources of power are shaped by the rules of the prevailing institutional logics. Institutional logics provide the rules that legitimate whether positional, relational, or economic factors form the basis of leadership power and authority in organizations (March and Olsen 1989). For example, in the context of the publishing industry, institutional logics indicate how the positions of founder, editor, publisher, president, manager, and chief executive officer (CEO) are valued and understood.

Second, institutional logics determine what issues to attend to in controlling and rewarding political behavior (March and Olsen 1976; Ocasio 1997). Logics provide the rules of the game that shape the cognition of social actors in organizations (Powell and DiMaggio 1991; DiMaggio 1997). Given ambiguity and cognitive limitations on executive decision making, organizations are limited in their ability to attend to all aspects of their environments (March and Olsen 1976; Simon 1997). Hence, organizational decision makers are constrained to focus their attention on a limited set of issues. Institutional logics comprise a set of implicit rules of the game that regulate which issues, strategic contingencies, or problems become important in the political struggle among actors in organizations (Ocasio 1997). For example, whether power is allocated to those enhancing the prestige of the publishing house or to those improving market position depends on whether the prevailing institutional logic in the industry focuses attention on prestige or, alternatively, on market competition.

Third, the assumptions, values, beliefs, and rules that comprise institutional logics determine what answers and solutions are available and appropriate in controlling economic and political activity in organizations (March and Olsen 1976). For example, institutional logics regulate whether to control or reward publishing executives' focus on particular solutions, such as organic (i.e., internal) or acquisition growth strategies, building personal imprints, or developing market channels. Moreover, the prevailing institutional logics determine the likelihood and appropriate-

ness of leadership succession itself as a political behavior and a routine solution to the problems of organizational performance. Consequently, the utilization of executive succession as a solution to organizational problems of market instability depends not only on prevailing economic conditions but also on whether existing institutional logics legitimate succession as an appropriate response to market instability.

RESEARCH DESIGN

To examine the consequences of institutional logics on executive power and succession requires identifying the prevailing logics in the industry and specifying whether and how these logics may have changed over time. This research strategy is in the empirical tradition of Fligstein (1985, 1987, 1990), Edelman (1990), Fligstein and Brantley (1992), Sutton and Dobbin (1996), Ramirez, Soysal, and Shanahan (1997), Zhou, Tuma, and Moen (1997), and Ruef and Scott (1998), a tradition in which differences in the effects of independent variables across time periods are used to draw inferences about changes in institutional logics and conceptions of control across these same periods. This research design assumes that institutional logics cannot be directly measured through any one variable or set of variables. Instead, our methodology follows those of the authors cited above in connecting the quantitative research to a historical analysis of the prevailing logics.

Our methods improve upon prior research by the use of in-depth interviews with industry principals to define the institutional logics and to identify the historical time periods in which they prevailed. Our selection of the higher education publishing industry for the analysis allows us to document and analyze the historical shift from a professional to a market logic. However, an important limitation of this historical design and sampling method is that firms in which the logic of the professions is prevalent are likely to be privately held and do not publicly disclose performance data, a potentially important determinant of executive succession. This limits our ability to measure firm performance as is common in conventional studies of succession using publicly held firms (Ocasio 1994; Zajac and Westphal 1996).

Our method follows Doty and Glick (1994), who illustrate the use of typologies in theory building and modeling. To identify the institutional logics and how they changed, we conducted taped interviews with higher education publishers, investment bankers who specialize in publishing, and staff of the Association of American Publishers (AAP). These interviews are part of a larger research project on higher education publishing. In 1991, 30 individuals were selected for interviews on the basis of their past and present leadership positions in the industry. These interviews

807

were transcribed and used to supplement historical sources and to formulate and ground the theoretical hypotheses. The appendix includes further information on the interviews. In addition, we conducted historical research using the industry trade literature, the publisher case histories written by the historian John Tebbel, and other books and articles on publishing (see quantitative data and methods section). From this information, we developed a theoretical typology, an abstract model of two ideal types of institutional logics: an editorial logic, which prevailed during the 1960s and early 1970s, and a market logic, prevailing since the mid-1970s. Each logic represents a combination of attributes that are hypothesized to be the determinants of executive power and succession in organizations.

## FROM AN EDITORIAL TO A MARKET LOGIC

According to our interviews and historical analysis, publishers described the 1950s and 1960s in higher education publishing as characterized mostly by small houses that were privately owned by families and persons who engaged in publishing as a lifestyle and a profession. The dominant form of leadership was the founder-editor, whose legitimacy and authority stemmed from their personal reputation in the field, their position in the organizational hierarchy, their relational networks with authors, and the stature of their books (Coser, Kadushin, and Powell 1982). The founder-editor's expertise was embodied in the individual person, and because of the uncertainty over the precise ingredients of a best seller, these leaders were accorded professional status (Hirsch 1972). Table 1 summarizes the characteristics of the two ideal types of institutional logics.[3]

During this era, publishers viewed their mission as building the prestige and the sales of the publishing house. To do so, they focused their attention on strategies of organic growth, hiring and developing editors with the best reputations to build personal imprints, develop new titles, refine the backlist of existing titles, and nurture relationships with authors (Asser 1989). Capital was committed to the firm for the longer term, and the leader's life cycle and family estate plans were the salient determinants of executive succession. We refer to this first set of ideal type attributes as the *editorial logic*.

The prevalence of an editorial logic during this time is exemplified by comments from the executive vice president in charge of strategic planning for a major higher education publisher:

[3] We have validated the two institutional logics with the phenomenological experience of executives in higher education publishing. As evidence that these logics have been found to be consistent with executives' own understanding, these ideal types are currently used at a well-known university publishers' college in their executive leadership training program.

TABLE 1

Two Ideal Types of Higher Education Publishing

| | Editorial Logic | Market Logic |
|---|---|---|
| Characteristics ................. .. | Personal capitalism | Market capitalism |
| Organizational identity . _ . | Publishing as profession | Publishing as business |
| Legitimacy .. .. .. .. .. ........ | Personal reputation | Market position |
| | Rank in hierarchy | Rank in performance |
| Authority structures .......... .. | Founder-editor | CEO |
| | Personal networks | Corporate parent firm |
| | Private ownership | Public ownership |
| Mission .. .. .. .. .. ........ .. | Build prestige of house | Build competitive position |
| | Increase sales | Increase profits |
| Focus of attention ............. .. | Author-editor networks | Resource competition |
| Strategy .................. ... .... . | Organic growth | Acquisition growth |
| | Build personal imprints | Build market channels |
| Logics of investment . .. ...... | Capital committed to firm | Capital committed to market |
| Rules of succession ............. . | Family estate plans | Market for corporate control |

> In the 1960s, publishing was a different world. Most of the companies were small and private. Nobody talked about profits; sales, yes, but not profits. . . . A lot of the publishing companies in those days were still run by the grand old men of publishing. I used to see Mr. Knopf come in every day with his white hair and his cane and walk into his dark blue velvet office with a great mahogany desk. There were truly devoted editors, who were really into literature. . . . And so this world was really not about business, and nobody cared that much about making a lot of money. You went into publishing because you liked authors and books.

In another interview, a former president and CEO in the early 1980s of one of the largest companies with both trade and higher education divisions talked about the historical change in the salience of personal reputation and relational networks with authors. He said, "when Prentice Hall bought Allyn and Bacon from the family in 1952, we asked about the royalty rate paid to authors. I remember how incredulous I felt when I heard the answer—they said it depends on whether they had a good year or not."

Within the period of an editorial logic, there were companies that operated as hierarchies—for example, the larger companies such as Prentice Hall, McGraw-Hill, and Macmillan. In addition, some of the venerable old-line publishers, such as Wiley and Harcourt Brace, became hierarchies in the 1960s (Moore 1982; Morris 1994). When William Jovanovich became president of Harcourt in 1960, he took the company public and began to mold it into a diversified hierarchy. However, at the same time, he

continued to run the publishing interests from an editorial logic, centered around a dominant individual with he himself editing manuscripts (Tebbel 1981). The growth of publishing hierarchies added the attribute of rank in the hierarchy as a salient characteristic of organizational identity under an editorial logic.

Based on our interviews and historical research, publishers described a shift that occurred in the organizational identity of higher education publishing sometime during the 1970s—a shift from publishing as a profession to publishing as a business. With this change, the dominant form of leadership became the CEO, whose legitimacy and authority stemmed from the firm's market position and performance rank, the corporate parent firm, and public shareholders. The mission was to build the competitive position of the firm and increase profit margins. To do so, the focus of executives' attention changed to counteracting problems of resource competition by using strategies such as acquisition growth and building market channels. This attention to "marketing" books is in sharp contrast to the older editorial logic where it was believed that good books sold themselves by favorable word of mouth (Powell 1985, p. 10). Hence, there was little point in investing in marketing a good book—people either have or lack the capacity to appreciate genius (Lane and Booth 1970, p. 42). Tebbel (*Microsoft Encarta 97 Encyclopedia* [CD-ROM], *s.v.* "book trade") reinforces this point by noting that in the 1960s modern marketing methods were rare in publishing. However, by the early 1980s, most publishers were emphasizing the most advanced marketing techniques. The logic of investment is to commit capital to its highest market return, and the salient rules of succession are shaped by the market for corporate control. We refer to this second set of ideal type attributes as the *market logic*.

One veteran publisher summarized this new market logic as follows:

> If you take it back to the 1960s, I remember seeing some things that were odd by publishing standards at the time. . . . The conglomerate phenomenon was one. It was not only the big companies outside the industry buying publishers, but there were some internal examples. . . . What sticks in my mind was the guy who put together InText. Buying up all those little companies to make one big important company. We real publishers looked at this and wondered—why was he doing this? This didn't fit publishing as we knew it. . . . All of a sudden what were really editors were now managers. The outside conglomerates gave up and divested . . . they couldn't understand the business . . . that we don't break even until nine months into the year. . . . But the conglomerate acquisitions gave publishers a first glance at finance skills and a new business—investment banking. . . . Maybe that is why we now [1991] have a market for publishing companies. . . . Of course, market pressures now create a whole new problem for executive stability.

Another executive publisher described the heightened attention to marketing and to building market channels and how in his company the editorial focus became contested:

> In the early 1970s, when I was the executive in charge of a division, the company CEO had a serious discussion with me about how I had to get rid of all these little books. Even though my books were important in their fields and selling well, they were in small markets and required the same amount of a sales rep's time—time that could be spent selling a book for a larger market. . . . But my real recognition of how this business had changed came when the parent company asked us not for editorial talent but for management talent for their other divisions. It was the realization that our mission was to grow managers, not book editors, that really shook me.

We also found support for a rise in market logic in the publishing industry literature. For example, Greco (1996, 1997) describes what he termed a "substantive reconfiguration" within publishing attributed to the direct impact of strategic planning practices on executives. Shatzkin (1982) comments on the commercialization of publishing strategies and its impact on the declining prominence of the editorial function. In reference to scholarly publishing, Powell (1985, p. 12) described a "shift in power within publishing houses—one in which editors are in decline and corporate managers and marketing are in ascendance." With respect to college text, scholarly, and trade publishing, Coser et al. (1982, p. 29) noted "a shift in the internal status order within publishing houses—a process in which the power of editors declined . . . and the influence of professional managers has risen." Last, the publishing historian Tebbel (1987, pp. 463, 464) also describes the decline in the influence of editors and the rise of market influences on publishing. He notes, "When the giant conglomerates stretched their tentacles into the book business, the moves sometimes brought into the publishing world a kind of executive not seen before." In another book, Tebbel (1981, p. 511) continues to discuss this point by saying, "Management was now in the hands of business-oriented people, while those who had combined business with editorial creativity were out of control."

A number of factors contributed to the decline of an editorial logic and the rise of a market logic. Haveman and Rao (in press) argue that when "segregating" processes—such as changes in competition, new political processes, the atrophy of social networks, new views of legitimacy, and new technologies—occur, they create pressures that contradict the prevailing logic and give rise to a new one. A review of the industry literature and time series data indicates that several of these processes occurred. In the early 1970s, there began a period of transition in logics, which was

propelled by new sources of capital in the industry, an increase in resource competition in the product market after the mid-1970s, new sources of information from trade presses that emphasized a focus on market logics, and the development of investment banking practices and firms specialized to the industry.

The antecedents for the changes in logics were evidenced by changes in market demand and the need for new sources of capital. In the 1960s, market demand exploded along with the demographic expansion of postwar baby boomers en route to college and with increased state and federal investments in the construction of new colleges and universities (Coser et al. 1982; Brint and Karabel 1991). Figure 1 shows the continuous increase in college enrollments prior to 1975, with a tapering off of the rate of increase afterward. Similarly, the sales of college-level books, approximately $67 million in 1956, had grown to more than $531 million in 1975, indicating that publishers responded to the increased demand in the product market (*Bowker Annual of Library and Book Trade Information* [1958, 1975]). With this growth, Wall Street analysts began to tout higher education publishing as a growth industry, signaling to corporate executives outside the industry, who were engaged in the heralded diversification strategies of this time (Fligstein 1990), that publishing firms were attractive targets for acquisition (Powell 1980; Coser et al. 1982, p. 25). Faced with both market growth and increasing competition, publishers needed new sources of capital (Smith 1995). As a result, family-estate publishers faced two choices: going public to obtain access to public capital markets or securing funding by being acquired.[4]

The increase in demand led to an increased number of publishing organizations and a change in the level of resource competition (see fig. 2). After 1975, resource competition in the product market became a salient issue because of the decline in the rate of increase in college enrollments (see figure 1), the acquisitions campaign into the American marketplace by foreign conglomerates (Graham 1994; Levin 1996), and the entry of nontraditional competitors specializing in course packs and used books (Bernstein Research 1994; Baker and Hileman 1987).

For publishing companies that were acquired, one consequence was that they became divisions and subsidiaries of corporate parent firms. Par-

[4] Acquisitions in the publishing industry should not be confused with hostile takeovers (Davis and Stout 1992). In examining all *Literary Market Place* and *Publishers Weekly* reports of acquisitions for the observation period, there were only three hostile takeover attempts in the 1980s, all of which failed. Furthermore, no hostile takeovers in the industry were found using SearchBank. When publishers were queried in interviews, they corroborated these findings that hostile takeovers are rare. As the president of the higher education division of one of the largest publishers put it, "Why would we have hostile takeovers? The assets in this business walk out the door every night."
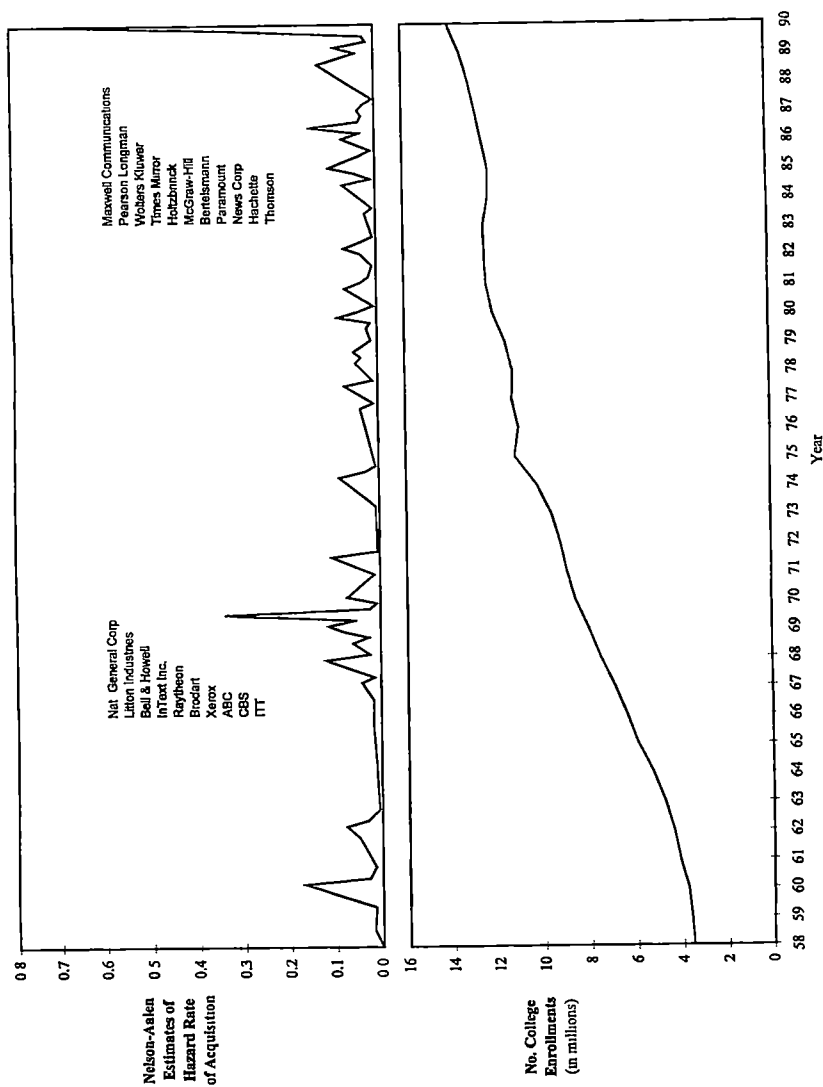
FIG. 1.—Industry time-series data for hazard rate of acquisition and number of college enrollments for higher education publishing, 1958–1990.
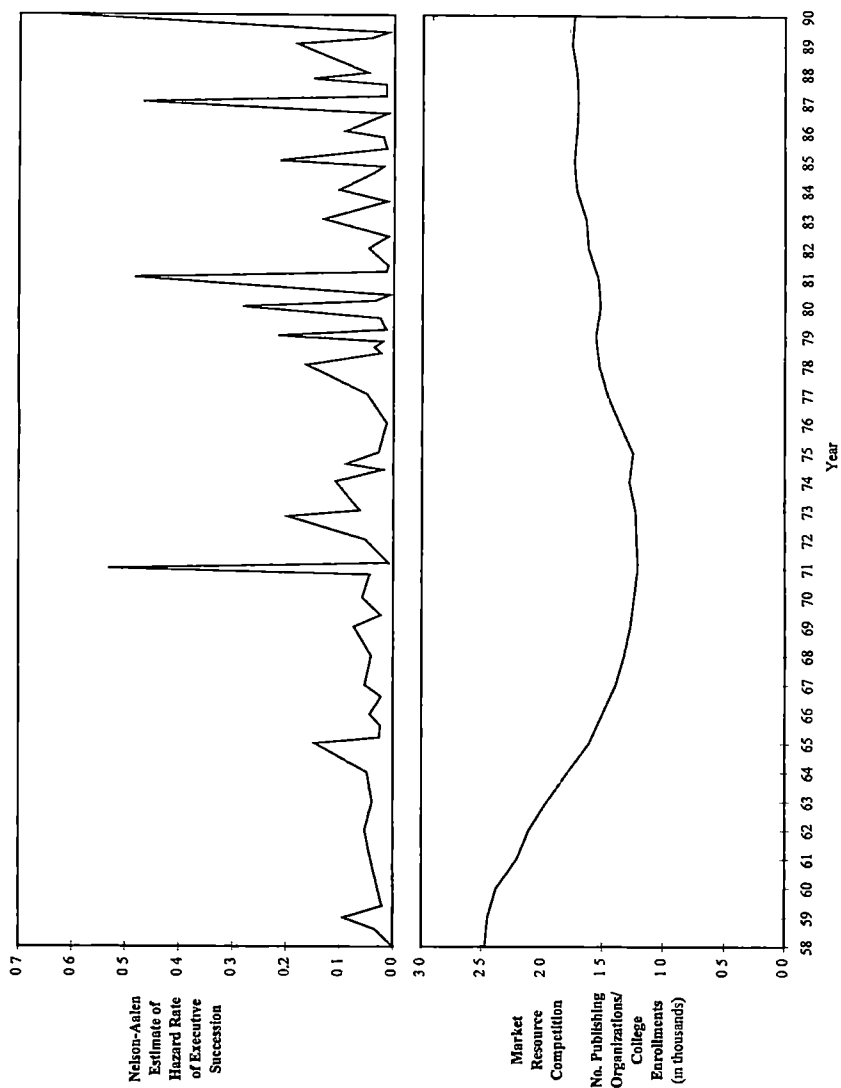
FIG. 2.—Industry time-series data for hazard rate of succession and market resource competition for higher education publishing, 1958–1990.

ent corporations superimposed on these publishers new performance expectations for yearly increases in profits and market share. This in turn refocused executives' logics of investment on market processes and on a new solution—the strategy of acquisition growth. For example, one publisher stated, "Instead of being able to manage your business for the value of future cash flow, you had to manage it for yearly profits transferred to the parent company. . . . Every year had to be better than the previous year. The only way to get bigger rapidly is to go outside and acquire others. Then you set up a new kind of industry competitiveness, which is: I want to buy this other company because if I don't our competitors will get it. So the attention shifts from publishing to what it is we can buy."

Executives told us that market position and reputation, which had previously taken years to establish under the editorial logic, could be obtained overnight with acquisitions. Similarly, the analysis by the communication scholar Greco (1989) reinforces this point. Moreover, for both public and private firms, attention to strategies of acquisition growth and the market for publishing companies created new determinants of executive succession by changing the sources of power and the rules for tenure in the position. As one executive said, "We were competing with rival divisions in other companies and had to overbid in order to say to our parent company that we won the property and that we were better. In hindsight, the later deals did not make good economic sense, so in order to make the acquisitions pay, we were going to have to plan on a lot of consolidation. This displaced a lot of people."

The institutionalization of the market logic was further evidenced with the founding of the *BP Report on the Business of Book Publishing* in 1977, a trade newsletter that targeted subscriptions to the executive suite (Abrahamson and Fombrun 1994). Rather than the typical *Publishers Weekly* (PW) features about new books, authors, and imprints, this newsletter focused on competitive position, ranking publishers by their control of market share, and providing information on acquisition practices as a means to increase market share. "Acquiring parent," "target company," and "deal price" were terms used for the first time in the publishing trade literature. One of the most basic indices of cultural centrality is the structure of language itself (Zucker 1983, p. 33; Hirsch 1986). This "linguistic framing" of market concepts increased their salience in the minds of publishing executives.

The institutionalization of the market logic was further propelled by the development of investment bankers specialized to publishing. Interviewees indicated that during the first acquisition wave in the late 1960s "deal makers" were neither former publishers nor tightly connected to the industry. Then, those deal makers came from Wall Street, and the acquiring firms were located in industries outside of publishing. However, in

the market period, as one CEO stated, "Investment bankers are now wired into the process." This practice has continued to formalize and to legitimate acquisition growth as a strategy to accomplish the mission of building the competitive position of the firm. Investment bankers now conduct training for publishers in how to "stay ahead of the game" by using acquisitions and consolidation as a business strategy (Fulcrum Information Services 1998, p. 2).

In contrasting the two industry-level logics in table 1, we note evidence of how editorial and market logics are identified with and shaped by societal-level logics of the professions and markets, respectively (Friedland and Alford 1991). The professions are organized bodies of experts who have authority and autonomy because they legitimate social missions (Freidson 1986) and because they have the ability to apply esoteric knowledge to particular cases (Abbott 1988, pp. 99–100). As indicated by our interviews and the industry literature, the traditional ideology of the industry was founded on viewing publishing as a profession.[5] The logic of the professions permeates the editorial logic, with editorial reputation and author-editor networks as key foci of attention in implementing the mission and strategy of the firm. However, a closer examination of the editorial logic also reveals the impact of another societal logic: the logic of the corporation. Both rank in the organizational hierarchy and the personal reputation of editors are key sources of legitimacy in the industry. The logic of the corporation is also revealed in the commitment of capital to the firm, rather than to the individual editors. Publishing houses under the editorial logic were perhaps best described as quasi-professional firms, where the ideology of the profession is intermixed with a formal hierarchy. Unlike pure professional practice firms—such as traditional law, public accounting, and architecture firms—publishing firms required no formal certification for gaining entrance to the profession, one reason publishing came to be labeled "the accidental profession" (Coser et al. 1982, p. 100). With the shift to a market logic, the professional orientation of the publishing industry declined and was replaced by the logic of Wall Street investment bankers and the increasing concern with profitability and market orientation common to other U.S. industries (Davis and Stout 1992; Useem 1996).

[5] The printer's mark, the origin of a publisher's imprint, symbolizes publishing as a profession (e.g., in the early history of publishing, the Tree of Knowledge of the Estiennes, and in the 1960s and 1970s, the motto Education for Truth of Wadsworth). The functional backgrounds of the founders of higher education publishers also evidence the connection to a profession. To name a few, George Bacon of Allyn and Bacon was a high school principal, Richard Prentice Ettinger of Prentice Hall was a college professor, and John Wiley was active in supporting church missionary efforts (Tebbel 1972, p. 269; Wiley 1999).

HYPOTHESES

The two institutional logics—editorial and market—make salient different potential sources of power and provide different rationales for whether and when to use executive succession as a solution to the issues confronted by organizations. Based on our interviews and theory, we expect these differences to translate into different effects for the independent variables in the two periods.

Positional and Relational Determinants of Succession

Research indicates that the position of founder is associated with lower rates of executive succession (McEachern 1975; Ocasio 1994). McEachern (1975) argues that founders have lower rates of succession because they have greater economic and political power relative to other executives. Carroll (1984) notes that founders have persistence in their positions because they have personal characteristics that distinguish them from non-founders—they are more likely to be owners, to have higher commitment to the organization, and to possess special expertise and knowledge. Ocasio (1999) found that founders have lower rates of succession because founder-led firms have not experienced succession and therefore lack organizational-level rules and routines that guide executive succession. Because of the lack of rules, succession is less likely to be an available solution to the problems of the organization.

According to our theory, both the legitimacy of potential sources of executive power and the salience of the rules of succession are likely to be modified by the prevailing institutional logics. First, the attributes of the ideal types in table 1 indicate that the personal and positional sources of power have greater legitimacy during the period of an editorial logic than during the period of a market logic. This implies that the effects of founders, which embody the attributes of personal and positional power, will be greater under an editorial logic. Second, the appropriateness of succession as a control mechanism for firms' actions and outcomes is likely to be greater in period 2 because attention to firm performance is more salient under a market logic. Extending this argument to the effects of founders, when a market logic prevails, firms without organization-level experience with succession are likely to rely on industry-level rules for succession as an organizational solution. The combination of these two factors leads to the expectation that the (negative) effect of founder on the rate of executive succession will be greater in the historical period when an editorial logic prevails.

Early research focused on the effect of bureaucratization, measured by organization size, on the frequency of executive succession (Grusky 1961). Grusky (1963) argued that bureaucratization increases the existence of

rules and routinization and therefore the likelihood of succession of executives. Comparing the largest 26 and the smallest 27 Fortune 500 companies, he found a positive relationship between organization size and executive succession. Reexamination of early studies reveals that findings are mixed and that there is not a simple and direct relationship between organization size and succession (Gordon and Becker 1964). In subsequent studies, Salancik and Pfeffer (1980), Allen and Panian (1982), and Harrison, Torres, and Kukalis (1988) found a significant positive relationship; Puffer and Weintrop (1991), Boeker (1992), and Ocasio (1994) found no effect of organization size on succession. Pfeffer and Salancik (1978) point out one reason for the mixed findings on organization size and succession is that size may be confounded with the relational and political dynamics of an organization. They argue that the larger the organization, the greater the number of departments and the greater the subunit basis of power, and therefore the greater the potential for contests for control.

According to our theory, the effects of organization size on succession may be conditional on whether the prevailing institutional logics focus executives' attention on the subunit basis of power associated with increased organizational size. A parallel measure of organization size and differentiation to the number of departments in an industrial firm is the number of imprints in a publishing firm; this measure is also an indicator of editorial control and editors' relational dynamics with authors. Publishing imprints represent a list of books that have identity and cohesiveness determined by an editor's professional expertise and author networks. According to the attributes of the ideal types in table 1, under an editorial logic, the focus of attention is on author-editor relationships. In such a climate, executives focused on the strategy of organic growth—ensuring the development of both new titles and a backlist of titles of personal imprints. Under a market logic, other subunits of the firm not controlled by editors—such as sales, marketing, and finance, which empower alternative strategies of growth, such as acquisitions and building market channels—should gain in importance. Therefore, the importance of imprints is likely to decline under a market logic where editor-author networks, inherently associated with the development of imprints, become less salient. Consequently, the (positive) effects of the number of imprints on the rate of executive succession will be greater in the historical period when an editorial logic prevails.

Structural position in the organizational hierarchy, derived from formal authority, is an important but relatively unexplored determinant of executive succession. This effect remains understudied because most empirical analyses of succession focus on executives at the same level of the organizational hierarchy, typically the CEO of independent firms. An exception is Boeker's (1992) study of semiconductor firms, which found that chief

executive officers were less apt to experience succession than lower ranked executives, who were more likely to be used as "scapegoats" when the performance of the firm was poor. In our study, we compare the succession rates of CEOs of division and subsidiary publishing firms with CEOs of independent publishers. According to table 1, we expect the effects of rank in the hierarchy to be moderated by the prevailing institutional logic. In particular, our interviews and historical analysis suggest that the legitimacy of rank and position in the organizational hierarchy as a determinant of executive power is likely to be greater in the period of an editorial logic than under a market logic. Given that divisional executives are of lower rank and status than executives of independent firms, we expect the (positive) effects of divisional executives on the rate of succession will be greater in the historical period when an editorial logic prevails.

Economic Determinants of Succession

According to our theory, with a change in the prevailing institutional logic from an editorial to a market focus, the determinants of executive succession are expected to shift from those based on positional and relational power to those based on economic power. We examine the effects of public versus private ownership as a determinant of succession. Research indicates that different forms of ownership imply different mechanisms for institutionalizing power in a firm (Pfeffer and Salancik 1978). McEachern (1975) found that owner-managed firms, and Allen and Panian (1982) found that family-owned firms, have increased executive tenure. However, Boeker (1992) found that public versus private ownership had no effect on executive succession. Given the increased salience of economic forces under the market logic and the shift from private to public ownership, as theorized in table 1, we expect the (positive) effects of public ownership on the rate of executive succession will be greater in the historical period when a market logic prevails.

Research indicates that changes in corporate control through acquisitions are likely to be followed by above-normal levels of executive succession in target firms (Walsh 1998). According to our theory and historical analysis of the development of the market for corporate control (Lazonick 1992), the effects of the market for corporate control should vary historically with a shift in the prevailing institutional logics. Moreover, our interviews and historical analysis of the industry indicate that the practice of acquiring firms to replace executives of target firms was more commonly accepted in the period when a market logic was dominant. Therefore, we expect the (positive) effects of acquisition on the rate of executive succession will be greater in the historical period when a market logic prevails.

Last, we examine the effects of product market competition on succes-

sion. Given data limitations in our sample, we cannot directly measure firm-level performance for most firms. We will evaluate instead how resource competition affects executive succession. Previous research on succession has not investigated the effects of competition on succession within a population of firms. According to our theory, the focus of executive attention shifted from author-editor networks to resource competition in the product market. Thus, performance in the product market is likely to gain salience in determining executive power in the period when a market logic prevails. Because of this, higher education publishing firms under a market logic will more likely view succession as a legitimate and appropriate remedy to issues of resource competition in the product market. Consequently, we expect that the (positive) effects of resource competition on the rate of executive succession will be greater in the historical period when a market logic prevailed.

## QUANTITATIVE DATA AND METHODS

The quantitative data set was constructed from archival data and from the results of a telephone survey on commercial higher education publishing firms. We drew upon the collections of the R. R. Bowker and the American Booksellers Association (ABA) libraries, which archive the most comprehensive sources of information on the publishing industry. We used as primary data sources four well-known publications: *Literary Market Place* (*LMP*), *Publishers Weekly* (*PW*), *Educational Marketeer* (*EM*), and *BP Report on the Business of Book Publishing* (*BP*).

*Industry and market definition.*—The book publishing industry consists of organizations that publish several types of books, such as trade, children's, school text, college text, scholarly, and professional reference books. At any given point in time, the higher education market is defined as those publishers that report in the *LMP* that they sell books in the college and university market.

Some publishers that sell books in the college and university market also publish books for other markets. During the observation period, some organizations in the sample moved into and out of different publishing product markets. On average, over the observation period, approximately 33% of the publishing firms in the sample published only one type of book, 40% published two types, 18% three types, and 9% four types of books. On average, approximately 8% of the firms in the sample publish only college texts, and 4% publish only scholarly books. Diversification also occurs within lists. As one editor who publishes for the college market stated, "We publish all manner of excellent books, including advanced texts, reference books and monographs" (Dougherty 1998). In 1976, books classified as general and mass market made up 34.8% of the higher educa-

tion market (Compaine 1978, p. 171). To test whether type of publisher and degree of market diversification might possibly affect our findings, we computed a 0/1 dummy variable, which was set equal to "1" if a firm in the sample published books for more than one type of market. We also computed a 0/1 dummy variable set equal to "1" if a firm in the sample published only textbooks. These variables were not statistically significant in any of the models, nor did they change any of the effects of the control and theoretical variables.

*Organization definition.*—The publishing organization is defined as either an independent firm or a division or subsidiary of a larger parent firm. For example, Wadsworth is a division of the Thomson Corporation, but it is counted as a separate case because it has a separate listing in the *LMP,* a separate organizational structure, a division president, and a different geographical location, distinguishable from the parent firm.

*Population and sample definition.*—The sample was randomly drawn from the population of all commercial publishing organizations listed in the *LMP* in any given year from 1958 to 1990 that reported publishing for the college and university market, a total of 766.[6] One-third of this population (230 publishing organizations) was selected as a sample, using the SPSS-X random-sample generator. First published in 1940 by a commercial publisher, the *LMP* remains the directory used industry wide by publishers, suppliers, distributors, writers, literary agents, bookstores, and librarians to identify whom to contact in conducting business in the publishing industry. Because the *LMP* lists the names, positions, and phone numbers of key personnel, the data in the *LMP* are kept current by the annual distribution of questionnaires to all organizations that publish a minimum of three titles a year.

The sample selection method used here differs from population-level studies of organizational founding and mortality that use the homogeneity of organizational form to define the population boundary (Hannan and Carroll 1992). Because organizational form is an independent variable and because institutional logics coalesce under conditions of status competition and social comparison among players in a product market (White 1981, 1992), we defined the population boundary according to a product market in which a variety of organizational forms cooperate and compete.

Organizations are in the sample from the date (year) that they first ap-

---

[6] Because the data prior to 1958 are excluded from the analysis, the observations are subject to left-truncation for firms in the sample founded prior to 1958. According to Tuma and Hannan (1984), this is not a serious problem because the piecewise exponential models used here lead to consistent estimates when both the age clock for each firm and the tenure clock for each executive are not restarted in 1958. The clocks instead begin with the firm founding date and executive first year of tenure.

pear in the *LMP* until the date that they are delisted because they no longer report information on the company and its employees. To ensure that our coding of delisted firms was accurate and not an artifact of a firm's missing one or two years' listings in the *LMP,* firms were traced through 1995, five years beyond the end of the observation period. One reason delisting occurs is because of business failure. With the exception of delisting because of acquisition, we treat delisting as a right censoring event, not as a succession event. We obtained information on acquired organizations because we theorize about how acquisition affects succession. Executive succession in the year subsequent to acquisition is coded as a succession event, whether or not the acquired firm was integrated into the parent firm or remained as a separate organization.

Dependent Variable

The dependent variable is the hazard rate of executive succession. We identified the executive as the top person listed in the *LMP* with line responsibility for the organization. To determine when executives were succeeded, executive names were compared across consecutive years for each organization for each year that the organization was listed in the *LMP.* Succession events are coded by year. Organizations can have multiple executives over the observation period.

We do not distinguish between "voluntary" or "forced" successions. Instead, following Puffer and Weintrop (1991) and Ocasio (1994), we control for retirement effects both by controlling for retirement age and by estimating models for a subsample that includes only executives 63 years old or less. The distinction between "voluntary" and "forced" successions is difficult to discriminate, given that executive turnover, even when not explicitly forced, typically entails a "push" factor associated with the loss of executive power. Political pressures on CEOs that result in executive succession need not imply a forced dismissal. CEOs may choose to depart and seek employment in other organizations when they have lost control over the firm's political coalition. Furthermore, research indicates that the information available for distinguishing between these categories is not reliable, even for publicly held firms (Beatty and Zajac 1987) and is almost nonexistent for privately held firms, which constitute approximately three-fourths of our sample. Moreover, our interviews with publishing executives indicated that the information that organizations release to employees and the business press is euphemized for a number of reasons. It is important to both executives and the board of directors to protect the firm from disruption, to maintain the market power of executives so that they may find alternative positions, to prevent lawsuits, and to ensure the payout of executives' financial incentive plans.

## Independent Variables

The variable *D* founder is a 0/1 dummy variable set equal to "1" in every year that the executive in office also was the founder of the organization. A chief executive listed in the *LMP* at the date of the company founding was assumed to be the founder. We used a telephone survey to ascertain the name of the founder when the chief executive name did not correspond.

We computed the variable ln *N* imprints, the natural logarithm of the number of publishing imprints per firm as listed in the *LMP,* as a measure of organization size and differentiation. We used the logarithmic transformation because the distribution of the number of imprints is skewed. An example of an imprint in higher education publishing would be an advanced mathematics series or a list of books on the sociology of culture. Other measures of organization size, such as the number of titles and the number of employees, are not reliable measures in the publishing industry (Coser et al. 1982, pp. 38–41). We also considered the use of firm assets as a measure of size. However, because the sampling procedure used here required temporal variation in forms of ownership (public and private) and organizational structure (independent and divisional or subsidiary), it was not possible to obtain financial data for all organizations in the sample for the entire observation period.[7] For similar reasons, we were unable to examine the effects of corporate profitability or those of the structure of the boards of directors.[8]

The variable *D* division/subsidiary is a 0/1 dummy variable set equal to "1" in every year that the organization was listed in the *LMP* as a division or subsidiary of a larger parent firm. The division/subsidiary variable was reset to "0" or to "1" if there was a state change in this variable in any given year.

We computed the variable *D* public/private ownership a 0/1 dummy variable set equal to "1" in every year that the subject organization is

[7] We considered collecting financial data on a subsample of publicly traded organizations to observe whether the same theoretical relationships found in the full sample would hold for publicly traded organizations. To explore this possibility, we analyzed the ownership characteristics of our sample. In 1958, 88% of the organizations in the risk set were independent; in 1975, 66%; and in 1990, 48%. In cross-tabulating independence (not being a division or subsidiary of a parent firm) and public and private ownership, we found that in 1958 only 3% of the organizations were both independent and public; in 1975, 4%; and in 1990, 2% Given the small numbers of independent, publicly held firms, there are insufficient observations to analyze the subsample.

[8] Although most contemporary studies of executive succession address these two factors, in our sample, these data could not be obtained for privately held firms nor for firms that were divisions of publicly held corporations.

publicly traded. Data on ownership were obtained from the Standard and Poors and Moody's Manual industry directories, *Ward's Business Directory of Major U.S. Private Companies,* and *Ward's Business Directory of U.S. Private and Public Companies.*

We computed the variable $D$ acquired a 0/1 dummy variable set equal to "1" if the subject organization was acquired. Acquired organizations were coded "1" only in the month and year of acquisition and "0" for every subsequent year in which acquisition did not occur. The acquisition variable was coded from all transactions listed in the *LMP* section on mergers and acquisitions. The *LMP* section on mergers and acquisitions also references the industry- and business-press articles that describe the transactions in detail. All these cited articles were read to verify the identity of the acquired and acquiring firms. In the case of succession due to acquisition, our interviews with industry principals indicate that there is generally a 3- to 6-month period where executives of acquired firms are given incentives to remain in their position in order to smooth the transition to the new owners.

To measure the effect of resource competition, we computed the ratio $N$ organizations/enrollment. This variable is a ratio measure composed of an annual count of potential competitors (all publishing organizations) and available resources (the number of college and university enrollments). College and university student enrollment data were obtained from the *Digest of Education Statistics,* published by the National Center for Education Statistics (1993). The resource competition variable is consistent with the ecological research that compares measures of population density to resource availability (Barron, West, and Hannan 1994).

## Control Variables

To control for executive retirement effects, we computed the variable executive age post-63. Executive age was coded in actual years and is transformed as the number of years over age 63. We conducted a telephone survey to obtain the ages of executives. Two organizations in the sample refused to provide information about the ages of their chief executives. During the observation period, 84 of the organizations in the sample were disbanded. We obtained age information for executives of disbanded organizations in one of two ways. If these individuals continued to be employed in the industry, we contacted them at the organizations in which they were subsequently employed. For retired or deceased individuals, five key leaders with long histories in the industry helped us identify missing age data. Using these methods, we were able to obtain the ages of over 60% of the chief executives of firms in the sample. We estimated

values for missing age data based on the mean value of actual age data. We computed a 0/1 flag variable $D$ estimated age set equal to "1" for age values that were estimated.

To control for the effects of tenure, executive tenure, we computed a cumulative count of the number of years an executive is in office (Ocasio 1994). Each executive begins with a value of "1," which is incremented by "1" for each additional year the executive remains in office.

We also controlled for succession effects due to organizational age by calculating the variable organization age from the organizations' founding dates as listed in the *LMP*. We used student enrollment data to control for demand fluctuations in the higher education market by computing a variable for percentage change in college enrollments, the percentage of change in student enrollments from the year $t - 1$ to the year $t$ divided by enrollments in the year $t - 1$.

Organizational form can be a central mechanism to embody and propel institutional logics (Haveman and Rao 1997). Fligstein's (1987) work showed that CEOs with marketing and finance backgrounds were favored in their rise to the top of multidivisional form organizations, M-form, but not in unitary form organizations, U-form. The argument is that power struggles and the sources of change in power are located in organizational structure because it locates resources available to actors: information and authority (Fligstein 1987, p. 46). To control for changes in the prevalence of organizational strategy and form, we compute two variables: the proportion of acquisition activity and the proportion of organizations that are divisions and subsidiaries of parent firms in the product market. The proportion of acquisition activity is the yearly hazard rate of acquisition of firms. The proportion of divisions and subsidiaries is the percentage of these types of organizations in the sample in any given year.

## Selection of Time Periods

Our interviews and literature reviews suggest that the transition between the editorial logic and the market logic occurred during the 1970s (Tebbel 1981; Shatzkin 1982; Coser et al. 1982; Powell 1985). Both of the Powell books captured this transition because the sample was drawn from the 1975 *LMP*, and the interview and survey data were collected from 1976 to 1978. To select the exact time periods for the hazard rate models, we rely on graph analysis of time series data of indicators associated with the changes in institutional logics. Furthermore, we undertake a sensitivity analysis to determine whether our results are sensitive to the specific cutoff period. Based on comparisons of these interview and literature review accounts, the graphical analyses, and the model specifications, we divided

observations for empirical analysis into two historical periods, 1958–75 and 1976–90, which correspond with the two ideal types of institutional logics shown in table 1.

Figure 1 shows graphs of two indicator variables: the Nelson-Aalen estimates of the hazard rate of acquisition and a market demand measure, the number of student enrollments in higher education for each year in the observation period. Figure 2 shows the hazard rate of the dependent variable, executive succession, and a measure of the covariate, resource competition, that is, the ratio of the number of publishing organizations divided by the number of student enrollments in higher education.

The measures of market demand, resource competition, and acquisition activity suggest 1975 as a cutoff point. Note in figure 1 that market demand, as measured by the number of student enrollments, begins to level off after 1975. Note in figure 2 that, prior to 1975, resource competition is decreasing, and after 1975, it begins to level and increase, signaling intensifying competition in the marketplace. With respect to the influences of acquisition activity, interviews and literature reviews established that there were two periods of acquisition activity, and the hazard plot in figure 1 supports these observations for our sample of firms (Thornton 1995). In the first wave, in the late 1960s and early 1970s, most acquired publishers were privately owned and were acquired by Fortune 500 firms, which were outside the industry and diversifying into higher growth markets (Powell 1980). In the second wave, in the 1980s and onward, acquirers were conglomerate publishers attempting to increase market share by horizontal integration under conditions of rising resource competition (Greco 1997). The hazard graph of acquisition in figure 1 identifies the main acquiring firms in the two time periods. The time series data on market demand and resource competition appear to be consistent with the characteristics of acquisition activity in the two time periods.

Last, we performed a sensitivity analysis to empirically examine the selection of the two time period cut points. Using an exponential model including all covariates, we estimated 22 time periods set at half-year intervals for three years before 1975 and eight years after 1975. While 1975 is the best fit, the statistically significant differences between the two time periods remain if we select the cut point at any time between 1972 and 1983.[9] This indicates a transition period during which cut points are not

---

[9] The sensitivity analysis suggests that logics may have changed incrementally between 1972 and 1983. Incremental institutional change is likely if change involves the process of hybridization, where organizations eventually replace some features of their current logics with one or more other logics (Zucker 1983; Haveman and Rao, in press). However, the objective of our analysis is not to determine whether the transformation was discontinuous or incremental but to test whether a transformation in logics of control affected the determinants of succession. Note also that an alternative model-

very sensitive. Effects change only if we select cut points before 1972 or after 1983. This is quantitative support for two distinct periods of institutional logics, which manifest their most significant effects earlier and later in the observation period.

### Selection of Models

We use event history analysis to examine the hazard rate of executive succession (Tuma and Hannan 1984). Our arguments about change in the effects of institutional logics require models that allow for effects to vary across time periods and levels of analysis—to vary with characteristics of individuals at risk of a succession event and with the characteristics of organizations and their environments. For this reason, piecewise exponential models are used, which allow the intercepts and the effects of covariates to vary in an unconstrained way across historical (calendar) time periods. The equation below summarizes the functional form of this model:

$$\log r_{j_{k_p}}(t) = \beta'_{j_{k_p}} x(t),$$

where the subscript $p$ denotes a given historical period (e.g., 1958–75, 1976–90), and $x(t)$ refers to the set of explanatory and control variables used in the analysis. Models were estimated by the method of maximum likelihood (ML) using RATE (Tuma 1993).

### RESULTS

Tables 2 and 3 present the mean values and the correlation coefficients for all variables in the models. Table 4 presents the ML estimates of piecewise exponential models on the rate of executive succession. To test for time variation in the "effects" of the variables, we compared two effects models: a time-invariant model (model A) and a time-varying model (model B). Model A has time-varying intercepts, which are listed under the historical period for which they apply, and time-invariant effects of covariates, which are listed under the column labeled "all years." The pattern of effects in the all-years model is analogous to an average effect for the 1958–90 period. For model B, time-varying parameter estimates are located under the historical period for which they apply. The coefficients give the effect of the covariate on the log of the rate of executive succession. A relative comparison of the scale of effects can be determined by taking

---

ing strategy of excluding a transition period could not be implemented, because it significantly reduces the sample size and limits the power of the test.

TABLE 2

MEANS, PEARSON CORRELATIONS, AND $F$-TESTS FOR COVARIATES OF HIGHER EDUCATION PUBLISHING FIRMS

| | TIME PERIOD | | | |
|---|---|---|---|---|
| | 1958–75 | 1976–90 | | |
| COVARIATE MEANS | Editorial | Market | $F$-VALUE | $p$-LEVEL |
| 1  Executive succession .............. ... .. ..... ..... | .05 | .06 | 1.59 | .207 |
| 2  Executive age post-63 .......... ... . ..... .... . | .35 | 96 | 56 59 | .000 |
| 3. $D$ estimated age  ...... ..... .. ... ...... ... . ..... | .35 | .34 | .83 | .361 |
| 4. Executive tenure . ... .. .. .. ...... ..... ..... .... | 10.25 | 11 05 | 6.77 | .009 |
| 5. Organization age  . ...... ...... .................. ...... | 34 13 | 28.37 | 24.97 | 000 |
| 6  % change in college enrollments .... .. ... .... | 6 92 | 1.53 | 6,086 15 | .000 |
| 7. Industry proportion division/subsidiary .. | .21 | 29 | 2,366 47 | .000 |
| 8  Industry acquistion activity . ..... ..... ..... | .02 | .04 | 197.99 | .000 |
| 9  $D$ founder ..... .... ... ...... . .... .... . ..... . ... | 55 | .54 | .48 | .487 |
| 10. ln $N$ imprints . ..... .... ... ... ..... ... . . . .. | .15 | 26 | 55 30 | 000 |
| 11. $D$ division/subsidiary . ..... .. ...... ..... ..... . | .28 | .35 | 21.36 | .000 |
| 12  $D$ public/private ownership .. .... ...... .... . | .22 | .27 | 9.40 | .002 |
| 13. $D$ acquired .... .. .......... . ....... ..... ..... . | .02 | 03 | 1 05 | .305 |
| 14. Resource competition ... .. ...... ..... ..... ..... | 1 58 | 1.63 | 32 85 | 000 |

NOTE —$N$ = number, $D$ = dummy variable.

the antilogs of the coefficients to provide the multiplier of the base rate. Two-tailed tests are used to interpret significance levels for the models as a whole. One-tailed tests are used to compare whether individual parameter estimates are significantly different between the two time periods, because our hypotheses are unidirectional.

The nesting of model A within model B allows the use of a likelihood ratio $\chi^2$ statistic to test the comparative fit of model A with model B. The likelihood ratio $\chi^2$ test comparing model A against model B indicates that model B, with both time-varying intercepts and effects, significantly improves the fit of the model ($\chi^2 = 27.33$; $df = 13$; $P \leq .01$). Thus, we can reject the model of time-invariant effects on the rate of executive succession when all covariates are considered simultaneously.

On the whole, the results support our overall hypothesis that with a shift from an editorial to a market logic, the determinants of succession changed from a basis of positional and relational authority to a basis of authority derived from economic and market forces. A $\chi^2$ contrast shows that the parameter estimates for the theoretical variables as an entire group cannot be constrained to be equal across the two time periods without significantly degrading the fit of the model ($\chi^2 = 17.27$; $df = 6$; $P \leq .01$).

TABLE 3

CORRELATION MATRIX FOR COVARIATES OF HIGHER EDUCATION PUBLISHING FIRMS

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | .03* | ... | | | | | | | | | | | |
| 3 | .00 | -.12*** | | | | | | | | | | | |
| 4 | -.02 | .51*** | -.12*** | | | | | | | | | | |
| 5 | .04* | .19*** | -.18*** | .14*** | | | | | | | | | |
| 6 | -.02 | -.08*** | .00 | -.02 | .07*** | | | | | | | | |
| 7 | .05** | .15*** | -.03* | .06*** | -.05** | -.50*** | | | | | | | |
| 8 | .05** | .09*** | -.04** | .05** | .03 | -.02 | .43*** | | | | | | |
| 9 | -.11*** | .09*** | .10*** | .36*** | -.47*** | .00 | -.04** | -.06*** | | | | | |
| 10 | .07*** | .03 | -.12*** | -.06*** | .09*** | -.09*** | .18*** | .08*** | -.20*** | | | | |
| 11 | .12*** | -.06*** | -.05** | -.13*** | .16*** | -.06*** | -.12*** | .05*** | -.34*** | .10*** | | | |
| 12 | .14*** | -.06*** | -.14*** | -.13*** | .24*** | -.03* | .00 | .05*** | -.37*** | .23*** | .50*** | | |
| 13 | .09*** | .01 | .02 | .01 | .02 | -.01 | -.01 | .03* | -.05** | .02 | .17*** | .13*** | .. |
| 14 | -.02 | .00 | -.03 | .01 | .04** | -.06*** | -.06*** | .11*** | -.06*** | -.03 | -.06*** | -.06*** | .01 |

NOTE — Heading numbers correspond to covariates listed in table 2.
* $P \le .05$, two-tailed tests
** $P \le .01$.
*** $P \le .001$.

TABLE 4

MAXIMUM-LIKELIHOOD ESTIMATES OF PIECEWISE EXPONENTIAL MODELS OF THE
RATE OF EXECUTIVE SUCCESSION

| | | TIME PERIOD | |
|---|---|---|---|
| COVARIATES | ALL YEARS | 1958–75 Editorial | 1976–90 Market |
| Intercept (model A)[a] | | −3.036*** | −3.435*** |
| Intercept (model B)[b] | | −3 402 | −4.646*** |
| Control variables: | | | |
| Executive age post-63 | .068** | .144* | .075** |
| D estimated age | .226 | .423+ | .116 |
| Executive tenure | .018* | .046*** | −.000 |
| Organization age | −.003+ | −.005 | −.003 |
| Pct chg coll. enrollments | −.060* | −.030 | −.084* |
| Industry proportion div /sub | .013 | −.002 | −.014 |
| Industry acquistion activity | 4 087 | 5.596 | 4.403 |
| Theoretical variables:[c] | | | |
| D founder | −.870*** | −1.125*** | −.650** |
| ln N imprints | .301** | .701** | .227+ |
| D division/subsidiary | .364* | 1.027*** | .175 |
| D public/private ownership | .671*** | .422 | .714*** |
| D acquired | .534 | −.855 | .945** |
| Resource competition | −.012 | −.083 | 1 334** |

NOTE —$N$ of firms = 230; $N$ of succession events = 237, $N$ = number; $D$ = dummy variable Likelihood ratio $\chi^2$ for model A = 172.57*** ($df$ = 26), for model B = 145.24*** ($df$ = 13); for difference between models = 27 33** ($df$ = 13)

[a] Model A has time-varying intercepts and time-invariant effects of covariates. Time-varying intercepts are shown under the historical period for which they apply. Time-invariant effects of covariates are listed under "all years."

[b] Model B has time-varying intercepts and time-varying effects of covariates Time-varying effects of covariates are shown under the period for which they apply.

[c] Model B theoretical variables cannot be constrained to be equal across historical periods without significantly degrading the fit of model B ($\chi^2$ = 17.27, $df$ 6, $P$ = 01) Constraining the control variables equal across periods marginally degrades the fit of the model ($\chi^2$ = 11 92, $df$ 7; $P$ = 10)

+ $P \leq .10$, two-tailed tests.
* $P \leq 05$
** $P \leq 01$
*** $P \leq 001$

With respect to the positional and relational basis of power, first we find that the effect of founders lowers the rate of executive succession. While this effect is as hypothesized, greater under an editorial logic than under a market logic, note that the $\chi^2$ contrast of the effect of founder between the periods is not strong, and we cannot reject the null hypothesis of no statistically significant difference ($\chi^2$ = 1.55; $df$ = 1; $P \leq .11$).

Second, increased organization size, as measured by the number of publishing imprints per firm, significantly increases the rate of executive suc-

cession in the editorial period ($P \leq .01$). The effect of organization size also is marginally significant in the period of a market logic. This effect on the base rate of executive succession is more than one and one-half times higher in the period of an editorial logic than in the period of a market logic ($.701 - .227 = .474$) ($e^{.474} = 1.61$). We found support for our hypothesis of a difference in the effects of number of imprints between the two periods. A $\chi^2$ contrast shows that the parameter estimates cannot be constrained to be equal across the two time periods without significantly degrading the fit of the model ($\chi^2 = 2.69$; $df = 1$; $P \leq .05$).

Third, we find support for our expectation about position in the hierarchy, that is, that executives in charge of a division or subsidiary of a parent firm are at a significantly higher risk of succession in the period of an editorial logic, but not in the period of a market logic. This variable is significant in the editorial period, at $P \leq .000$ and not significant in the market period. This effect on the base rate of succession is more than two times higher in the period of an editorial logic than in the period of a market logic ($1.027 - .175 = .852$) ($e^{.852} = 2.34$). Moreover, the parameter estimates for this positional effect cannot be constrained to be equal across the two time periods without significantly degrading the fit of the model ($\chi^2 = 6.03$; $df = 1$; at $P \leq .01$).

With respect to the economic bases of power, first we find that the effects of form of ownership, public versus private, are equivocal. The size and significance of the coefficients in the two periods is consistent with our expectations—not significant under the period of an editorial logic but positive and significant under the period of a market logic. However, note that the contrast in the effect of ownership between the periods is not strong, and we cannot reject the null hypothesis of no statistically significant difference ($\chi^2 = .68$; $df = 1$; $P \leq .20$).

Second, the effect of acquisition of the firm on the rate of executive succession is consistent with a market logic, emphasizing growth by acquisition. The parameter estimates are positive and significant for the period of a market logic but not for the period of an editorial logic. Consistent with our hypothesis, the two-parameter estimates for acquisition cannot be constrained to be equal across periods without significantly degrading the fit of the model ($\chi^2 = 4.02$; $df = 1$; $P \leq .02$). The effect of acquisition on the base rate of succession is strong, being six times higher in the period of a market logic than in the period of an editorial logic ($.945 - (-).855 = 1.800$) ($e^{1.800} = 6.05$).

Third, the findings support our expectation that the fate of executives was tied to the issues of scarcer resources and higher competition in the product market during the period of a market logic. The parameter estimate for the effect of resource competition is positive and significant during the period of a market logic but not during the period of an editorial

logic. The effect of resource competition on the base rate of executive succession is four times higher in the period of a market logic than in the period of an editorial logic $(1.334 - (-).083 = 1.417) (e^{.417} = 4.12)$. Consistent with our hypothesis of there being a greater effect of resource competition under a market logic, the estimates for resource competition cannot be constrained to be equal across the two time periods without significantly degrading the fit of the model $(\chi^2 = 2.74; df = 1; P \leq .05)$. This finding, while suggestive, is subject to potential bias given that firm-level performance data are unavailable for the study.

## Control Variables

The estimate for executive age is positive and significant in both time periods, indicating that, independent of the effects of different institutional logics, executives due to retire because of age have a higher rate of succession. The estimate for the effect of executive tenure is positive and significant in the period of an editorial logic and negative and not significant in a period of market logic. A positive and significant tenure effect suggests that increasing tenure may lead to obsolescent strategies that in turn trigger adversity among political coalitions in the firm and executive succession (Ocasio 1994). This finding also may suggest one mechanism by which an editorial logic was displaced with a market logic over time.

The control variable for market demand, percentage change in student enrollments, is negative and significant in the all-years model, indicating that, on average, greater demand in the product market decreases succession. Note this effect is significant in the period of a market logic but not in the period of an editorial logic. This period difference provides some support for our arguments that during the period of a market logic, market forces influenced attention to the political processes of leadership succession. The parameter estimates for the proportion of firms that are divisions and subsidiaries of parent firms and the rate of acquisition activity in the industry are not significant in either the all-years or the piecewise models. While graph analyses of these variables show that their proportion is increasing over time in the sample, the results from our models suggest that the effects of a change in institutional logics on executive succession was not due to these industry-level changes in strategy and organizational form.

## Subsample Analysis: Succession Prior to Retirement Age

To further examine if our results may be affected by forced versus voluntary succession, we control for voluntary retirement (Puffer and Weintrop 1991; Ocasio 1994) by estimating separate models for the effects of succes-

TABLE 5

Maximum-Likelihood Estimates of Piecewise Exponential Models of the Rate of Executive Succession for the Subsample, Executives under Age 63

| COVARIATES | ALL YEARS | TIME PERIOD 1958–75 Editorial | TIME PERIOD 1976–90 Market |
|---|---|---|---|
| Intercept (model A)[a] | | 3.447*** | −3.805*** |
| Intercept (model B)[b] | | 2.146 | 4 761*** |
| Control variables: | | | |
| D estimated age | .260+ | 458+ | .110 |
| Executive tenure | .021+ | .039* | .002 |
| Organization age | −.004 | − 004 | −.003 |
| % change in college enrollments | − 063* | −.042 | −.083+ |
| Industry proportion division/subsidiary | 023 | .028 | − 011 |
| Industry acquistion activity | 3.280 | 2.513 | 4.251 |
| Theoretical variables:[c] | | | |
| D founder | − 947*** | −1.063*** | −.786** |
| ln N imprints | 202 | .596* | 116 |
| D division/subsidiary | .407* | .972** | .212 |
| D public/private ownership | .570*** | 506+ | .536** |
| D acquired | .658+ | −.602 | 1.093** |
| Resource competition | 065 | − 471 | 1 228* |

Note.—$N$ of firms = 226, $N$ of succession events = 199, $N$ = number, $D$ = dummy variable. Likelihood ratio $\chi^2$ for model A = 139 01*** ($df$ = 24), for model B = 120 65*** ($df$ = 12); for difference between models = 18 36+ ($df$ = 12)

[a] Model A has time-varying intercepts and time-invariant effects of covariates. Time-varying intercepts are listed under the historical period for which they apply Time-invariant effects of covariates are listed under "all years."

[b] Model B has time-varying intercepts and time-varying effects of covariates Time-varying effects of covariates are listed under the period for which they apply

[c] Model B theoretical variables cannot be constrained to be equal across time periods without significantly degrading the fit of model B ($\chi^2$ = 14 74; $df$ 6; $P$ = 02) Control variables can be constrained to be equal across periods ($\chi^2$ = 4 48; $df$ 6; $P$ = 61)

+ $P \leq$ 10, two-tailed tests
* $P \leq$ 05.
** $P \leq$ 01.
*** $P \leq$ 001

sion of executives prior to retirement age (63 years or younger). Note that in our sample approximately 85% of the successions are for executives 63 years or younger. As shown in table 5, the results of this diagnostic test are consistent with our theory and show that our statistically significant effects of the theoretical variables on the rate of succession hold for executives prior to retirement age. A $\chi^2$ contrast shows that the parameter estimates for the theoretical variables as an entire group cannot be constrained to be equal across the two time periods without significantly degrading the fit of the model ($\chi^2$ = 14.74; $df$ = 6; $P$ = .02).

More specifically for this subsample, the positional and the relational bases of power, as measured by the number of organizational imprints and divisional versus independent executives, significantly increased the rate of succession in the period of an editorial logic but not in the period of a market logic. The economic bases of power as measured by acquisition of the firm and resource competition significantly increased the rate of executive succession in the period of a market logic but not in the period of an editorial logic. The $\chi^2$ contrast of the effects of these variables between the two periods is strong, and we can reject the null hypothesis of no statistically significant differences. The effects of founder and public versus private ownership, while in the expected direction, are not strong enough to reject the null hypothesis of no statistically significant differences between the two periods.

## DISCUSSION AND CONCLUSION

We extend prior research on executive succession by providing a new set of findings on the historical contingency of the determinants of executive power and succession. The interviews and historical research established that the prevailing institutional logic shifted in the 1970s from an editorial to a market focus. The event history models suggest that this historical change in logics led to different determinants of executive succession. The findings support our general argument—that when, whether, and how executives deploy their power to affect succession in organizations is conditional on the prevailing institutional logic in an industry.

In particular, the findings suggest how the salience of positional, relational, and economic determinants of power varies by historical time period. The quantitative analysis implies that changes in logics result not in an overall increase in the rate of leadership succession, but in countervailing determinants of succession in the two periods. With the transition from an editorial logic to a market logic, the effects on the rate of succession of organizational size and position in the hierarchy declined, while those of acquisition and resource competition increased.

The effects of organization size and rank and position in the hierarchy were significantly stronger in the period when an editorial logic prevailed, even though as indicated in tables 2 and 3, organization size and the proportion of multidivisional organizations were significantly lower in this period. In the market period, neither organization size nor the executive's position as division head or independent CEO were strong predictors of succession. Organization size was a strong predictor in the editorial period when executives focused their attention on strategies of organic (internal) growth. However, size loses significance when internal growth strategies waned in favor of an alternative strategy in the market period—growth

by acquisition. The effects of divisionalization are particularly revealing. Under an editorial logic, the power of independent CEOs relative to division heads led to higher rates of succession for the latter group. Under a market logic, rank and position in the hierarchy is less important, so that the difference in rates of succession between CEOs at the top of the hierarchy and for divisions is not statistically significant.

With respect to economic determinants, the effects for firm acquisition and resource competition were significantly stronger when a market logic became dominant, even though acquisitions and resource competition were noteworthy during both periods. Both our historical analysis and interviews shed light on how to interpret these findings. The publishing historian John Tebbel (1981, p. 724) noted, "it was true that publishers had been surviving since the beginning of the nineteenth century through mergers and acquisitions—but the changes were mostly shifts in partnerships and within the family, so to speak." Tebbel's historical observation suggests that in period 1, when "family governance" of acquisitions was commonplace, executives were not displaced during acquisition as they were in period 2, when the "market governance" of acquisitions became commonplace. Moreover, we went back to a few of the publishers that we interviewed to discuss these findings. They indicated that many of the acquisitions in the late 1960s were made by product-unrelated conglomerate firms outside the industry. These acquiring firms had neither the incentive nor the structural means to consolidate publishers into existing operations. Instead, the acquiring firms were dependent on extant publishing leaders to run their acquired properties as relatively autonomous divisions—allowing for the prevalence of an editorial logic even upon acquisition. In contrast, the second acquisition wave, in the late 1980s, took place among product-related publishing conglomerates in a period when a market logic had taken hold. This made executives susceptible to a market logic, where attention was focused on improving market position by strategies of acquisition and integrating operations—suggesting why acquisition led to succession.

Resource competition had no effect on succession during the editorial period, although the level of resource competition—the number of publishing organizations per thousands of college enrollment—was actually higher in the early 1960s than at any time during the period when the market logic prevailed. This finding may suggest that attention to economic forces depends on the prevalence of a market logic rather than on the level of competition experienced in the industry.

The findings support our overall hypothesis that the sources of power that affect executive succession are not universal constants but are historically contingent. However, for two of our original six hypotheses, the null hypothesis of historically invariant effects could not be rejected. These

null findings suggest that some of the determinants of executive power and succession may be less historically contingent than others. For example, we found lower rates of succession for founders in both periods. This might be explained in two ways. First, the staying power of founders may be due to their personal characteristics, such as having higher commitment and expertise, which may be more stable across time and situation than is true of organizational and environmental characteristics (Scott 1995, p. 137). Second, it may also be that founder-led firms, which lack experience with succession, do not rely on industry-level rules, but they instead rely on organization-level rules (Ocasio 1999).

A null finding was also uncovered for the differential effects of public and private ownership between the two periods. The strength of the main effect indicates that some sources of power are less historically contingent than others and reveals a limitation or scope condition of our theory. Note in table 4 the effect of public ownership is statistically significant in the market period but not in the editorial period. However, the contrast between the findings in tables 4 and 5 indicates that this apparent historical contingency may be due to a retirement effect, not to involuntary succession. The lack of historical contingency for ownership form suggests that any effects of the rise of a market logic affected both privately and publicly held firms, even though privately held firms are thought to be less susceptible to direct market pressures.

The results of both the interviews and quantitative analysis are consistent with the view that institutional logics are both material and symbolic (Friedland and Alford 1991). A shift from an editorial logic to a market logic in higher education publishing was marked by an increase in the size of publishing organizations, public ownership, and resource competition—all structural characteristics consistent with the increased importance of market forces in the industry. These structural changes in market conditions attracted new and powerful actors with different goals and tactics that comparatively de-emphasized intrinsic editorial accomplishment and elevated financial pursuit. However, after controlling for structural and economic forces at the industry level, the effects of our theoretical variables remain. This suggests that without an accompanying change in the understandings that comprise an institutional logic, economic and structural changes may not be sufficient to explain the determinants of executive succession.

Generally, we found that there is not a one-to-one correspondence between changes in structural and economic forces and changes in executive power and succession. The changing determinants of executive succession held independently of the control variables changes in student enrollments, the proportion of multidivisional firms, and the rate of acquisition activity in the industry. This evidence suggests that institutional logics

moderate the effects of economic and structural forces affecting succession. This implies that the effects of institutional logics cannot be reduced to purely social structural and economic forces (Barley and Kunda 1992; DiMaggio 1994). This evidence supports Tilly's (1997, p. 78) and Stinchcombe's (1987) arguments that history should be taken seriously in the study of social structures and that shared understandings and their object-ifications constrain social interactions (Fligstein 1996).

Our interpretation of the findings is not meant to imply that logics and meanings are completely independent of changes in the social structure or in the economy. Institutional logics must articulate with prevailing structural conditions in an industry (Wuthnow 1989), providing a set of meanings, interpretations, and symbols that make sense of material reality (Barley and Kunda 1992). In our example, the rise of a market logic in higher education publishing—articulated with observed changes in public ownership, acquisition activity, and resource competition—allowed publishers to understand these changes and develop suitable responses. The editorial logic—with its emphasis on publishing as a profession rather than a business, its emphasis on author-editor networks, and its emphasis on personal reputation and rank in the hierarchy—could not readily explain or account for the changes in the marketplace nor the rise of acquisition activity after 1975.

The findings suggest that changes in institutional logics, while serving to articulate changes in economic and social structures, are not epiphenomenal but lead to changes in the determinants of executive power and succession. We propose the view that the relative autonomy of institutional logics operates through the way logics structure the attention of organizational decision makers. The findings suggest that institutional logics, once they become dominant, affect succession by structuring the attention of executives toward the set of issues that are consistent with the logic dominant within an industry, whether editorial or market, and away from issues that are not. While changes in key variables—such as organizational size, position in the hierarchy, acquisitions, and resource competition—may occur independently of institutional logics, whether these changes are attended to and whether or not they are consequential for decisions on executive power or succession is contingent on whether the prevailing logic makes these variables salient.

The higher education publishing industry provides evidence of the historical contingency of executive power. Future research should examine whether and how these effects change if firm-level performance measures are available. Other data limitations, particularly on firm size and CEO age may also affect the interpretation of our results. While the specific findings of this study may not be generalizable to other industries, the theory that we test on the historical contingency of power in organizations

provides a novel approach that is generalizable across different settings. Future research should investigate how power and succession are affected in other industrial settings that have experienced significant transformation. One example for future study is the health care sector, which has also experienced a transformation away from professional logics (Starr 1982) toward managerial and market logics (Ruef and Scott 1998). We also see parallel transformations with privately held firms in accounting, consulting, investment banking, and law—shifting from a professional logic to a market logic. In addition, the furniture and funeral home industries, originally craft based, were subjected to rationalization in the 1980s by the market for corporate control and therefore are interesting sites for comparative industry studies of the consequences of institutional change on power in organizations.

The rise of the market logic in higher education publishing and in other professional and craft industries parallels higher-order transformations in the United States (Useem 1996). Since the 1970s, managerial capitalism in the United States has been increasingly subject to pressures of the marketplace, the financial community, and the market for corporate control (Jensen 1993). The owners of capital have gained increasing control over corporate managers. While these macrolevel changes have been observed primarily in settings traditionally dominated by corporate logics, such as Fortune 500 industrial firms (Davis and Stout 1992), parallel changes are observed in higher education publishing, although there the transition developed from a professional logic to the dominance of a market logic.

The historical contingency of power in organizations and the decline of the professions relative to markets as mechanisms through which power is constituted has significant implications for the development of products such as books and health care that have cultural and political significance. More generally, the transition of institutional logics from professions to markets implies that a different set of values determines the production of products and the distribution of resources in organizations and in society.

## APPENDIX

### Interview Methodology

Publishers representing experience at both the editorial and executive levels in organizations of varying age, size, and structure were contacted by telephone and asked to suggest individuals they believed had broad experience and important reputations in the industry during the observation period. Publishers identified by their peers were invited to be interviewed. The chief executives in this sample often began their careers in entry-level sales and had worked in both sales and editorial positions for several publishing organizations.

A 23-question interview was administered to publishers from Boston, New York, and San Francisco. We asked publishers to describe how the higher education publishing industry changed from the 1960s to the end of the 1980s with respect to leadership, management strategy, market structure, products, and technology. Many of the questions were open-ended and required recollections of past events. Respondents were asked to put themselves in the frame of reference of the particular historical time and to try not to use the benefit of hindsight. Similarly, a 38-question interview was administered to three investment bankers who were identified by publishers as representing the key banking firms to the industry. In addition, two directors of well-known university presses were interviewed, one of whom was the president of the Association of American University Presses. Interview protocols were approved by the Human Research Subjects Committee at Stanford University. All respondents except one agreed to have their interviews tape recorded.

## REFERENCES

Abbott, Andrew. 1988 *The System of Professions: An Essay on the Division of Expert Labor.* Chicago: University of Chicago Press.

Abrahamson, Eric, and Charles J. Fombrun. 1994. "Macrocultures: Determinants and Consequences." *Academy of Management Review* 19 (4): 728–55.

Allen, Michael P., and Sharon K. Panian. 1982. "Power, Performance and Succession in the Large Corporation." *Administrative Science Quarterly* 27:538–47

Asser, Paul Nijhoff. 1989. "Consolidation, Internationalization, and the Future of Publishing: A Scenario." *Book Research Quarterly* 5 (Fall): 51–59.

Baker, DeWitt C., and James Hileman. 1987. "Used Books and the College Textbook Industry." *Book Research Quarterly* 3 (Fall). 8–17.

Barley, Stephen R., and Gideon Kunda. 1992. "Design and Devotion: Surges of Rational and Normative Ideologies of Control in Managerial Discourse." *Administrative Science Quarterly* 37:363–99.

Barron, David N., Elizabeth West, and Michael T. Hannan. 1994. "A Time to Grow and a Time to Die: Growth and Mortality of Credit Unions in New York City, 1914–1990." *American Journal of Sociology* 100 (2): 381–421.

Beatty, Randall P., and Edward J. Zajac. 1987. "CEO Change and Firm Performance in Large Corporations: Succession Effects and Manager Effects " *Strategic Management Journal* 8·305–17.

Bernstein Research. 1994. *The School Publishing Industry.* New York: Sanford C. Bernstein.

Boeker, Warren. 1992. "Power and Managerial Dismissal: Scapegoating at the Top." *Administrative Science Quarterly* 37:400–21.

*Bowker Annual of Library and Book Trade Information,* sponsored by the Council of National Library Associations. 1958–95. New York: R. R. Bowker.

Brint, Steven, and Jerome Karabel. 1991. "Institutional Origins and Transformations: The Case of American Community Colleges." Pp. 337–60 in *The New Institutionalism in Organizational Analysis,* edited by Walter W. Powell and Paul J. DiMaggio. Chicago: University of Chicago Press.

Carroll, Glenn R. 1984. "Dynamics of Publisher Succession in Newspaper Organizations." *Administrative Science Quarterly* 29:93–113.

Compaine, Benjamin M. 1978. *The Book Industry in Transition: An Economic Study of Book Distribution and Marketing.* White Plains, N.Y.: Knowledge Industry Publications.

Coser, Lewis A., Charles Kadushin, and Walter W. Powell. 1982. *Books: The Culture and Commerce of Publishing.* Chicago: University of Chicago Press.

Davis, Gerald F., and Henrich R. Greve. 1997. "Corporate Elite Networks and Governance Changes in the 1980s." *American Journal of Sociology* 103 (1): 1–37.

Davis, Gerald F., and Suzanne K. Stout. 1992 "Organization Theory and the Market for Corporate Control: A Dynamic Analysis of the Characteristics of Large Takeover Targets, 1980–1990." *Administrative Science Quarterly* 37:605–33.

DiMaggio, Paul 1994. "Culture and Economy." Pp. 27–57 in *Handbook of Economic Sociology,* edited by Neil J. Smelser and Richard Swedberg. Princeton, N.J.. Princeton University Press.

———. 1997. "Culture and Cognition." *Annual Review of Sociology* 23:263–87.

Doty, D. Harold, and William H. Glick 1994. "Typologies as a Unique Form of Theory Building Toward Improved Understanding and Modeling." *Academy of Management Review* 19 (2): 230–51.

Dougherty, Peter. 1998. "A Random Walk down Third Avenue." Paper presented at the annual meeting of the American Sociological Association, San Francisco, August.

Edelman, Lauren B. 1990 "Legal Environments and Organizational Governance: The Expansion of Due Process in the American Workplace." *American Journal of Sociology* 95 (6). 1401–40.

Fligstein, Neil. 1985. "The Spread of the Multidivisional Form among Large Firms, 1919–1979." *American Sociological Review* 50:377–91.

———. 1987. "The Interorganizational Power Struggle: The Rise of Finance Personnel to Top Leadership in Large Corporations, 1919–1979." *American Sociological Review* 52:44–58

———. 1990. *The Transformation of Corporate Control.* Cambridge, Mass.: Harvard University Press.

———. 1996. "A Political-Cultural Approach to Market Institutions." *American Sociological Review* 61:656–73.

Fligstein, Neil, and Peter Brantley. 1992. "Bank Control, Owner Control, or Organizational Dynamics: Who Controls the Large Modern Corporation?" *American Journal of Sociology* 98 (2): 280–307.

Freidson, Eliot. 1986. *Professional Powers: A Study of the Institutionalization of Formal Knowledge.* Chicago: University of Chicago Press.

Friedland, Roger, and Robert R. Alford. 1991. "Bringing Society Back In: Symbols, Practices, and Institutional Contradictions." Pp. 232–63 in *The New Institutionalism in Organizational Analysis,* edited by Walter W. Powell and Paul J. DiMaggio. Chicago: University of Chicago Press.

Fulcrum Information Services. 1998. "Doing the Right Thing at the Right Time." Paper presented at the Second Annual Book Publishing Industry Mergers and Acquisitions Institute, New York, December.

Gordon, G., and S. Becker. 1964. "Organizational Size and Managerial Succession: A Reexamination." *American Journal of Sociology* 70:215–33.

Gouldner, Alvin W. 1954. *Patterns of Industrial Bureaucracy.* Glencoe, Ill.: Free Press.

Graham, Gordon. 1994. *As I Was Saying: Essays on the International Book Industry.* London: Hans Zell Publishers.

Greco, Albert N. 1989. "Mergers and Acquisitions in Publishing, 1984–1988: Some Public Policy Issues." *Book Research Quarterly* (Fall):25–43.

——— 1996. "Shaping the Future: Mergers, Acquisitions, and the U.S Publishing,

Communications, and Mass Media Industries, 1990–1995." *Publishing Research Quarterly* (Fall):5–15.

———. 1997. *The Publishing Industry*. Boston: Allyn and Bacon.

Grusky, Oscar. 1960 "Administrative Succession in Formal Organizations." *Social Forces* 39 (2): 105–15.

———. 1961. "Corporate Size, Bureaucratization, and Managerial Succession." *American Journal of Sociology* 67:263–69.

———. 1963. "Managerial Succession and Organizational Effectiveness." *American Journal of Sociology* 69:297–317

Harrison, Richard J., David L. Torres, and Sal Kukalis 1988. "The Changing of the Guard. Turnover and Structural Change in Top-Management Positions." *Administrative Science Quarterly* 33:211–32.

Hannan, Michael T., and Glenn R. Carroll. 1992. *Dynamics of Organizational Populations: Density, Legitimation, and Competition*. New York: Oxford University Press.

Haveman, Heather A., and Hayagreeva Rao. 1997. "Structuring a Theory of Moral Sentiments." *American Journal of Sociology* 102 (6): 1606–51.

———. In press. "Hybrid Forms and Institutional Change in the Early California Thrift Industry." In *Bending the Bars of the Iron Cage: Institutional Dynamics and Processes*, edited by Walter W. Powell and Daniel Jones. Chicago: University of Chicago Press.

Hirsch, Paul M 1972. "Processing Fads and Fashions: An Organization-Set Analysis of Cultural Industry Systems." *American Journal of Sociology* 77 (4): 639–59.

———. 1986. "From Ambushes to Golden Parachutes: Corporate Takeovers as an Instance of Cultural Framing and Institutional Integration." *American Journal of Sociology* 91 (4) 800–837.

Jackall, Robert. 1988. *Moral Mazes: The World of Corporate Managers*. New York: Oxford University Press.

Jensen, Michael C. 1993. "The Modern Industrial Revolution, Exit, and the Failure of Internal Control Systems." *Journal of Finance* 48:831–80.

Lane, Michael, and Jeremy Booth 1970 *Books and Publishers*. Lexington, Mass.: D.C. Heath.

Lazonick, William 1992. "Controlling the Market for Corporate Control: The Historical Significance of Managerial Capitalism." *Industrial and Corporate Change* 1 (3): 445–88.

Levin, Martin P. 1996. "The Positive Role of Large Corporations in U.S. Book Publishing." *Logos* 7:127–37.

*LMP (Literary Market Place)*. 1940–95. New York. R. R. Bowker.

March, James G., and Johan P Olsen. 1976. *Ambiguity and Choice in Organizations*. Bergen: Universitetsforlaget.

———. *Rediscovering Institutions: The Organizational Basis of Politics*. New York· Free Press

McEachern, William A. 1975. *Managerial Control and Performance*. Lexington, Mass.: Heath.

Meyer, John W., John Boli, George M Thomas, and Francisco O. Ramirez. 1997. "World Society and the Nation-State." *American Journal of Sociology* 103 (1): 144–81.

Moore, John Hammond. 1982. *Wiley: One Hundred and Seventy-Five Years of Publishing*. New York: John Wiley & Sons.

Morris, Christopher. 1994. *A History of Harcourt Brace*. New York· Harcourt Brace.

National Center for Education Statistics. 1993. *Digest of Educational Statistics*. Washington, D.C.: National Center for Education Statistics.

Ocasio, William. 1994. "Political Dynamics and the Circulation of Power: CEO Succession in U.S. Industrial Corporations, 1960–1990." *Administrative Science Quarterly* 39:285–312.

——— 1997. "Towards an Attention-Based View of the Firm." *Strategic Management Journal* 18:187–206.

———. 1999. "Institutionalized Action and Corporate Governance: The Reliance on Rules of CEO Succession." *Administrative Science Quarterly,* 44 (2): 384–416.

Ocasio, William, and Hyosun Kim. 1999. "The Circulation of Corporate Control. Selection of Functional Backgrounds of New CEOs in Large U.S. Manufacturing Firms, 1981–1992." *Administrative Science Quarterly* 4 (3): 532–62.

Perrow, Charles. 1986. *Complex Organizations.* New York: Random House.

Pfeffer, Jeffrey. 1981. *Power in Organizations.* Marshfield, Mass.: Pitman.

———. 1992. *Managing with Power: Politics and Influence in Organizations.* Boston: Harvard Business School Press.

Pfeffer, Jeffrey, and Gerald Salancik 1978. *The External Control of Organizations.* New York: Harper & Row.

Porac, Joseph F., Howard Thomas, Fiona Wilson, Douglas Paton, and Alaina Kanfer. 1995. "Rivalry and the Industry Model of Scottish Knitwear Producers." *Administrative Science Quarterly* 40:203–27.

Powell, Walter, W. 1980. "Competition versus Concentration in the Book Trade." *Journal of Communication* 30 (2): 89–97.

———. 1985. *Getting into Print: The Decision-Making Process in Scholarly Publishing.* Chicago University of Chicago Press.

———. 1991. "Expanding the Scope of Institutional Analysis." Pp. 183–203 in *The New Institutionalism in Organizational Analysis,* edited by Walter W. Powell and Paul J. DiMaggio. Chicago: University of Chicago Press.

Powell, Walter W., and Paul J. DiMaggio, eds. 1991. *The New Institutionalism in Organizational Analysis.* Chicago: University of Chicago Press.

Puffer, Sheila M., and John B. Weintrop. 1991. "Corporate Performance and CEO Turnover: The Role of Performance Expectations." *Administrative Science Quarterly* 36 (1)· 1–19.

Ramirez, Francisco O., Yasemin Soysal, and Suzanne Shanahan. 1997. "The Changing Logic of Political Citizenship: Cross-National Acquisition of Women's Suffrage Rights, 1890–1990." *American Sociological Review* 62:735–45.

Ruef, Martin, and W. Richard Scott. 1998. "A Multidimensional Model of Organizational Legitimacy: Hospital Survival in Changing Institutional Environments." *Administrative Science Quarterly* 43 (4): 877–904.

Salancik, Gerald R., and Jeffrey Pfeffer 1980. "Effects of Ownership on Executive Tenure in U.S. Corporations." *Academy of Management Journal* 23:653–64.

Scott, W. Richard. 1995. *Institutions and Organizations.* Newbury Park, Calif.: Sage.

Shatzkin, Leonard. 1982. *In Cold Type: Overcoming the Book Crisis.* Boston: Houghton Mifflin.

Simon, Herbert Alexander. (1947) 1997. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations,* 4th ed. New York: Free Press.

Smith, Joel. 1995. *Understanding the Media: A Sociology of Mass Communication.* Cresskill, N.J.: Hampton Press.

Starr, Paul. 1982 *The Social Transformation of American Medicine.* New York: Basic Books.

Stinchcombe, Arthur L. 1987. *Constructing Social Theories.* Chicago: University of Chicago Press

Strang, David, and John W. Meyer. 1994. "Institutional Conditions for Diffusion." Pp. 100–112 in *Institutional Environments and Organizations: Structural Complexity and Individualism,* edited by W. Richard Scott and John W. Meyer. Thousand Oaks, Calif.: Sage.

Sutton, John R., and Frank Dobbin. 1996. "The Two Faces of Governance: Responses

to Legal Uncertainty in U.S Firms, 1955 to 1985 " *American Sociological Review* 61:794–811.

Tebbel, John. 1972. *A History of Book Publishing in the United States,* vol. 1, *The Creation of an Industry, 1630–1865.* New York: R. R. Bowker.

———. 1981. *A History of Book Publishing in the United States,* vol. 4, *The Great Change 1940–1980.* New York: R. R. Bowker.

———. 1987. *Between Covers: The Rise and Transformation of Book Publishing in America.* New York: Oxford University Press.

Thornton, Patricia H. 1995 "Accounting for Acquisition Waves: Evidence from the U.S. College Publishing Industry " Pp. 199–225 in *The Institutional Construction of Organizations: International and Longitudinal Studies,* edited by W. Richard Scott and Soren Christensen. Thousand Oaks, Calif.: Sage.

Tilly, Charles. 1997. "History and Sociological Imagining." Pp. 67–82 in *Sociological Visions,* edited by K. Erickson. Lanham, Md. Rowman & Littlefield.

Tuma, Nancy Brandon. 1993. RATE software program, ver. 3.0A.

Tuma, Nancy Brandon, and Michael Hannan. 1984. *Social Dynamics: Models and Methods.* San Diego, Calif.: Academic Press.

Useem, Michael. 1996. *Investor Capitalism: How Money Managers are Changing the Face of Corporate America.* New York: Basic Books.

Walsh, James P. 1998. "Top Management Turnover Following Mergers and Acquisitions." *Strategic Management Journal* 9:173–83.

Weber, Max. (1922) 1978. *Economy and Society: An Outline of Interpretive Sociology,* edited by Guenther Roth and Claus Wittich. Berkeley: University of California Press.

White, Harrison C. 1981. "Where Do Markets Come From?" *American Journal of Sociology* 87·517–47.

———. 1992. *Identity and Control: A Structural Theory of Social Action.* Princeton, N J.: Princeton University Press.

Wiley, Peter 1999. E-mail correspondence with the author, June.

Wuthnow, Robert 1989. *Communities of Discourse: Ideology and Social Structure in the Reformation, the Enlightenment, and European Socialism.* Cambridge, Mass.: Harvard University Press.

Zajac, Edward J., and James D Westphal. 1996. "Who Shall Succeed? How CEO/ Board Preferences and Power Affect the Choice of New CEOs." *Academy of Management Journal* 39:64–90.

Zerubavel, Eviatar. 1997. *Social Mindscapes: An Invitation to Cognitive Sociology.* Cambridge, Mass. Harvard University Press.

Zhou, Xueguang, Nancy Brandon Tuma, and Phyllis Moen. 1997. "Institutional Change and Job-Shift Patterns in Urban China, 1949 to 1994." *American Sociological Review* 62:339–65.

Zucker, Lynne G 1983 "Organizations as Institutions." *Research in the Sociology of Organizations* 2:1–47.

# Book Reviews

*Nationalism Reframed: Nationhood and the National Question in the New Europe.* By Rogers Brubaker. New York: Cambridge University Press, 1996. Pp. xi+202. $54.95 (cloth); $16.95 (paper).

E. J. Hobsbawm
*University of London*

The present collection of occasionally overlapping studies confirms Brubaker's standing as the most impressive of the younger generation of the scholars in the, now vast, field of nationalism. At all events, he is the one whose merits are most readily appreciated by historians, since his analyses always rest on, or are documented by, comparative historical case studies. It consists of six essays. A chapter on "Rethinking Nationhood: Nation as an Institutionalized Form, Practical Category, Contingent Event" is followed by studies of the nationalist heritage inherited from the USSR by its successors, by three essays developing the theme of the interaction between "national minorities, the newly nationalizing states in which they live, and the external 'homelands' to which they belong." They are also overshadowed by the problem of the former USSR, so, in practice, is the last chapter on "Aftermaths of Empire and the Unmixing of Peoples," that is, ethnic homogenization by migration and terror.

Although the author's range of reference is worldwide in theory, in practice the studies are confined to central and eastern Europe, and, in effect, to the ethnopolitical consequences of the collapse of the great polyethnic European empires after the two world wars and after 1989: the Ottomans, the Habsburgs, the Russian empire whose fall was postponed for three-quarters of a century by the October Revolution and the theoretically monoethnic German empire, whose conquests made it briefly, and disastrously, polyethnic. More specifically, the book focuses on Russia. It is much more cursory, though excellent, on the explosive zone of the Balkans.

The focus is thus both geographically and historically specific. Even in Europe, topics that cannot be fitted into the pattern of the three great 20th-century breakdowns do not quite fit into the book's pattern—for instance, the neoseparatist ethnic nationalism that became a factor in the politics of old-established western "nation-states" such as Britain or even Switzerland, well before it began to devastate the communist region.

The restricted focus of these essays limits its practical rather than its theoretical interest. "The analytical task at hand," Brubaker argues, "is to think about nationalism without nations," that is, without supposing that the "nation" exists as any kind of real entity or substantial collectiv-

ity. Following Bourdieu, he sees it as "produced—or better . . . induced—by *political fields* of particular kinds," and these rather than the properties of collectivities govern its dynamics. To make sense of what has happened in post-communist Europe, all the analyst needs to know is that "Nation is widely, if unevenly, available and resonant as a category of social vision and division" in the modern world. This is certainly the case in Europe, and it may be argued that, like other Western innovations, it is still conquering the rest of the globe. How it got to be such a "category of practice" is not immediately relevant to his purpose, although he is plainly aware of the complex debates about the way it did so in the USSR.

Yet the availability—both practical and emotional—of the category "nation" in its classical Western form can no longer be taken for granted, even though the category "sovereign territorial state" remains and, *via* censuses with ethnic classifications and the conflicts between the governments and national sports teams, keeps it alive. Could the Indian BJP party, Hammas, or the Taliban be adequately described as a religiously tinted nationalism, like the Catholicism of the nationalist Irish or Poles? Ethnic diasporas are worldwide; "homeland nationalism" is not. As Brubaker himself notes, the United States is by its construction unsuited to "national minorities" of the European kind. His global skepticism about the "nation" in theory cannot conceal a Eurocentric perspective in practice.

Nevertheless, his theoretical stance enables him to focus on certain important and neglected general aspects of the problem. The most urgent of these is "nationness as an event, as something that suddenly crystallizes," as tragically exemplified during the collapse of Yugoslavia and the USSR. He is equally perceptive about the peculiarities of the nationalism of "state peoples" in multinational societies, such as the Russians, or for that matter the English, for whom being Russian (or English) was only incidentally something that distinguished them from Finns or Scots, but primarily an overarching superethnic identity that was actually reinforced by the variety of intermarrying peoples who lived under a Tsar or Queen. He sees the asymmetry of nationalist discourse between the Russians and the other nationalities who defined themselves against Russians as a different people. (As Brubaker notes, the Russian language distinguishes, as the English language does not, between the territorial-political and the ethnolinguistic description of its people.) It is only today that Russians, English, and Spaniards are forced to think of themselves as "nations" in the same way as Poles, Scots, and Catalans. It is far from clear how they are to do so.

He also, incidentally, reminds us that, not least in the Balkans, "in the protracted course of postimperial migratory unmixings, the phases of greatest intensity have for the most part been closely linked to actual or threatened violence, especially during or immediately after wars." In other words, if you wish to see ethnic injustice transformed into mass expulsion and genocide, the best way to do so is to start a war. Alas, it is too late to heed this lesson in 1999.

*Revolutionary Demands: A Content Analysis of the* Cahiers de Doleances *of 1789.* By Gilbert Shapiro and John Markoff. Stanford, Calif.: Stanford University Press, 1998. Pp. xxxi+684. $75.00.

John A. Hall
*McGill University*

If the origins of this book go back to the mid-1950s when Gilbert Shapiro first began to teach on Tocqueville's interpretation of the revolution, it owes most to collaboration between the authors at the University of Pittsburgh over the last quarter of a century. This is a long time, but scarcely a second can have been wasted. For this is an awesome and pathbreaking achievement of which sociology as a whole should be proud: every angle of a key historical data source has been teased out and made available to systematic use, and in a way that for once really does immeasurably improve upon prior work of historians. All that a short review can do is to hint at some of the richness of very dense analyses: to that end, it makes sense to spell out the observation correctly made by Charles Tilly in his foreword, namely that there are in fact three different books within this large volume.

The first specifies the nature of content analysis in a high-powered and authoritative manner. The core of the argument made rests on a use of Chomskian linguistics to defend the need for human coders of data—as against those who imagine that content analysis could hereafter be done in absolutely aseptic manner by powerful computers. The extraordinary creativity of human language use means that mere raw data is, at least as yet, almost useless: what matters rather is the counting of statements endowed with meaning, a process which can only be achieved by human coders. Shapiro and Markoff have long experience with coders, and this leads to a mass of interesting insights. They found that non–social scientists coded their data most easily; the entry of data thus depends on generalists whose work can then be used by specialists. Further, they note that they are not really sure how such coders do their work, but they present and apply a series of ingenious tests to show how very reliable is the data so created. The whole of this methodological section is both exciting and convincing, and it deserves the widest readership: it makes a powerful case for the realiability and usefulness of a particular method and offers a model as to how a historical source can be turned into usable data for social science.

The second contribution is much more specific, and it concerns the nature of the *cahiers de doleances* (the lists of grievances) drawn up in 1789 for the consideration of the first Estates General to be called for 175 years. The *cahiers* were drawn up by members of the three great estates, a process which involved some 40,000 bodies, mostly rural. The authors are especially good at explaining, in the light of a large secondary literature, the nature of the evidence that the *cahiers* provide: they provide compelling arguments and tests that show that they are broadly rep-

resentative of the feelings of the different estates, noting in particular that the pressure of interlopers and of nobles was not such as to prevent popular feelings from being expressed. We learn most about the *cahiers*, however, by the exposition of the coding procedure. The complete coding scheme, printed as an appendix, makes it clear that a source has now been created that allows systematic analysis to replace mere impressionism, both in sheer numerical terms and in awareness of the complexities created by regional variation.

The final element at work here is a series of studies on the revolution itself. Some of these chapters have appeared previously, but they can be seen anew here as exemplars of some of the research techniques that can be applied to the data, from mere frequencies to sophisticated multivariate analyses. Still, the various studies—on the most common demands, sources of consensus and conflict between and within the three estates, the role of religion in the recalcitrant West, the relationship between social mobility and radicalism, the image of the King, and peasant grievances and insurrection—do make powerful contributions to understanding the revolution. For example, we are made aware of a process rather than a mere set of structures, are confirmed in the now current emphasis on a political account of discontent (but, importantly, without thereby undermining an appreciation of those socioeconomic differences that fueled the move from rebellion to revolution), and find Tocqueville's stress on the importance of the difference between *pays d'etats* (regions subject to centralized royal control) and *pays d'election* (possessed of relative autonomy from the state) strikingly confirmed. But perhaps the most striking lesson is that of the sheer complexity of the events codified as "the Revolution," and hence of the weakness of most extant explanatory accounts. In this context, one can express a hope, namely that the authors will one day try systematically to replace failed general accounts with a new one of their own—however spare it may prove to be.

*Histoire et sociologie en France: De l'histoire-science à la sociologie durkheimienne.* By Robert Leroux. Paris: Presses Universitaires de France, 1998. Pp. x+269.

Charles Tilly
*Columbia University*

Ideas resemble a river's floating debris: they come from many places and either wash up on shore, disintegrate, or drift out to sea, but for a while, they rush together in the same direction. By reflective observation of the river formed by French social thought since the mid-19th century, Robert Leroux has identified some surprising flows. Heavies of 19th-century historical theory such as Fustel de Coulanges, he argues, sought to build an autonomous, theoretically coherent history in reaction against both romanticism and grand philosophies of history. They eventually proved

incapable, however, of synthesizing the history of particular events with a history of universal causes. The "historical synthesis" identified with Henri Berr from the 1890s onward established an uneasy junction between sociological history and history proper. In the process, Émile Durkheim and his followers—here represented chiefly by Célestin Bouglé, François Simiand, and Maurice Halbwachs—then established a long hegemony over French definitions of appropriate relations between history and the social sciences. Durkheim's victorious vision portrayed sociology as a science of social evolution and history as the placement of particular peoples in their times and spaces. Even the famous *Annales* school of Lucien Febvre, Marc Bloch, and their brilliant successors worked under that Durkheimian hegemony, escaping the study of mere events in pursuit of something very much like Durkheim's program of historical sociology. Only today, as French historians (like such American counterparts as William H. Sewell, Jr.) call for a return to genuine events, do we see that hegemony cracking. So argues Leroux.

Let us situate that argument carefully. Leroux's materials come almost entirely from methodological and synthetic statements: prefaces, reviews, critiques, polemics, addresses, and the like. Except for some characterizations of Charles Seignobos's big summary volumes on French history, his analyses leave concrete historical practice almost untouched. He follows the history of ideas through general statements of those ideas rather than through their applications in historically oriented research programs or even their influence on prevailing interpretations of crucial historical events, processes, and entities. When, as a fledgling French historian, I hung around the *Annales,* the practitioners did (as Leroux's account implies) affect enthusiasm for theoretically informed history and disdain for mere political fact grubbing. In practice, however, no young historian gained recognition without a backbreaking accumulation of evidence and careful collation of that evidence with political history. Few rejected base-superstructure models, whether the base in question was successive modes of production, commercialization, changing demographic regimes, family structures, or the history of business cycles. Those few exceptions espoused the history of mentalities, which they insisted on documenting with the same dogged thoroughness as their cousins who reconstructed prices or birth rates. Historical practice, then, embodied models and doctrines only distantly connected with those of Durkheimian theorists. That is generally the case: although currently fashionable ideas circulate as easily through the sermons and prefaces of historians as of sociologists, historical practice generally changes significantly only when someone establishes a compelling new model joining a type of evidence, a form of analysis, and a style of argument.

As Leroux notes, François Simiand influenced historians far more than did any other Durkheimian. (Maurice Halbwachs served on the editorial board of the *Annales* but generally shunned historical analyses himself; his unfinished work on collective memory only came to historians' serious attention recently.) Although Simiand emitted prefaces, reviews, cri-

tiques, polemics, addresses, and the like with the best of them, he actually worked his influence through vast accumulations of time series, his interpretation of those time series in terms of business cycles, and his explanation of those business cycles by fluctuation in money supply. More than anyone else, he placed what he explicitly called the *longue durée* on the historical agenda. And in doing so, he broke with two crucial Durkheimian ideas: that simpler societies show us elementary forms of the same processes appearing in complex societies and that the general program of history should center on tracing the evolution of human societies.

Still, general doctrines do shape the founding of journals, the awarding of posts and prizes, the organization of meetings, and the presentation of a field's findings to the world at large. More so than their German and Anglo-Saxon counterparts, French historians and historically inclined sociologists have been disposed to declare their sympathies with a broadly evolutionary conception of social life. Robert Leroux's meticulous scholarship helps explain that predilection.

*Civility and Subversion: The Intellectual Dynamic in Democratic Society.* By Jeffrey C. Goldfarb. New York: Cambridge University Press, 1998. Pp. ix+253. $59.95 (cloth); $22.95 (paper).

Neil McLaughlin
*McMaster University*

We are all public intellectuals now. An exaggeration, of course, for sociology is and must remain primarily a theoretically driven and empirically grounded academic discipline. Yet Russell Jacoby's *The Last Intellectuals* (1987) certainly resonated with widespread dissatisfaction with the academic professionalism typified by both excessively technical statistical analysis and jargon-ridden postmodern theorizing. The founding of a new ASA journal for communicating with the public only reflects a growing belief that our discipline's proud tradition of social critics such as W. E. B. Du Bois, David Riesman, C. Wright Mills, and more recently, William Julius Wilson and Alan Wolfe must be brought closer to the center of professional sociology. Jeffrey Goldfarb's *Civility and Subversion: The Intellectual in Democratic Society* (1998) is a well-crafted manifesto for a renewal of democratic intellectual criticism and, as such, raises important questions for citizens, scholars, and sociologists.

Goldfarb's basic argument is straightforward and compelling. Modern society suffers from a "deliberation deficit." We live in an information-society culture bombarded with data and images created by spin doctors and ideologues, leaving little room for informed and rational debate on the major public moral, cultural, and political issues of the day. Arguing against the cultural cynicism he has written about elsewhere, Goldfarb makes a compelling case that intellectuals can play a vital democratic role in modern societies if they strike the right balance between civility

and subversion. Although hierarchy and judgement run against the grain of modern culture, intellectuals must promote cultural standards by distinguishing good from bad arguments and high-quality from mediocre ideas. Intellectuals must contribute to democratic and rational dialogue on issues that matter while disrupting common sense with sharp or even rude challenges to orthodoxies. The devil is in the balance, not so much as in the details.

*Civility and Subversion* is a theoretically informed analysis based on Goldfarb's own extensive experience as an observer and student of the democratic resistance to communism in Eastern Europe, as well as from observations about modern university life and race in the United States. Goldfarb illuminates his argument with carefully chosen and well-written discussions of such intellectuals as Walter Lippman, John Dewey, C. Wright Mills, Edward Said, Malcolm X, Toni Morrison, and Milan Kundera. This creative mix of scholarship and social criticism allows Goldfarb to link his analysis of the role of the intellectual to real world social movements and political debates. He has provocative things to say about "why there is no feminism after communism," illuminating both contemporary politics in Eastern Europe as well as how social movements face dilemmas as they cross cultural and national borders. Goldfarb also has thoughtful things to say about the polarized debate about political correctness on North American campuses, raising valuable critical questions about counterproductive forms of cultural radicalism, while insisting that excessive professionalism and careerism on the part of both students and faculty remains the major threat to a university life dedicated to ideas and democracy. Goldfarb's discussion of racial politics in America and the Thomas-Hill affair strikes a nice balance between disrupting racist discourse while encouraging engagement in level-headed dialogue about race in America. And finally, Goldfarb's thoughtful review of the literature on "what is an intellectual" combined with his extensive research on the democratic opposition to communism in Eastern Europe makes for a compelling case that the literature underestimates the importance of such cultural workers as artists, novelists, film makers, and musicians.

I would dissent from Goldfarb's argument in matters of tone, politics, and analysis, and these issues are linked. There is a self-referential tone to *Civility and Subversion*. Four "I"s in the opening paragraph, too much recounting of personal conversations with particular intellectuals, and an out-of-place airing of a minor professional conflict within the Sociology of Culture Section of the ASA get in the way of the larger issues the book addresses with such clarity and purpose.

Politically, Goldfarb is to be commended for raising critical questions about such left icons as Edward Said, C. Wright Mills, and Malcolm X while insisting on the vital role these subversives play in our political life. Yet Walter Lippman is one of the few conservative intellectuals Goldfarb discusses seriously as a role model for the civility side of his balancing act. While Goldfarb's knowledge of Eastern Europe politics provides an

invaluable comparative perspective, short case studies of such North American thinkers with complicated politics as David Riesman, Daniel Bell, Irving Louis Horowitz, and Seymour Martin Lipset would have added depth to a book that is really about left intellectuals.

The personal tone and the restricted politics of Goldfarb's book flows from the fact that *Civility and Subversion* is not really a scholarly analysis of intellectuals as it is a moderate manifesto based on some of our best sociology and the left liberal politics that many of us share. Just as intellectuals must find the right balance between civility and subversion, however, sociologists must balance our roles as scholars and researchers as well as citizens and intellectuals. Goldfarb largely avoids the romanticization of "public intellectuals" and the hostility to academia that mars Jacoby's work. And Goldfarb's analysis is a useful counterbalance to Randall Collin's tendency in *The Sociology of Philosophies* (1998) to privilege academic knowledge and undertheorize the relationship between scholars' and intellectuals' movements. Yet a sociology of intellectuals must be less personal, politically balanced, and more centrally concerned with theorizing the relationship between academic scholarship and popular and elite culture and politics. Ultimately, however, Goldfarb's *Civility and Subversion* raises all the right questions, is a good read, and will stimulate exciting research as well as professional and political activity.

*Outside the Fold: Conversion, Modernity, and Belief.* By Gauri Viswanathan. Princeton, N.J.: Princeton University Press, 1998. Pp. xx+332. $55.00 (cloth); $16.95 (paper).

Gerald James Larson
*Indiana University*

The thesis of Viswanathan's book involves three related assertions: (1) "conversion ranks among the most destabilizing activities in modern society" (p. xvi); (2) because it restores "belief from the margins of society to a more worldly function" (p. xvi) ("worldly" in the sense that "belief" comes to impinge on civil and political rights); and (3) that, therefore, "conversion is a subversion of secular power" (p. 3). The eight essays in the volume, at least four of which were published earlier as shorter pieces, are reworked here into eight "chapters" in three subdivisions: part 1, entitled "Dissent and the Nation" (including two essays or chapters), part 2, entitled "Colonial Interventions" (including four essays), and part 3, entitled "The Imagined Community" (including two essays). The book is primarily concerned with colonial India, but it also ranges further afield into an essay on John Henry Newman's conversion to Catholicism in 1845 (in chap. 2, entitled "A Grammar of Dissent") and to various observations about the Salman Rushdie affair with respect to the issues of "blasphemy" and "heresy" (in chap. 8, entitled "Epilogue: The Right to Belief").

The first chapter (entitled "Cross Currents") sets forth a theme that

recurs throughout the various essays, namely, that there is an interesting connection between, on the one hand, Macaulay's introduction of English education in India in 1835 (in order to produce Indians who would be "Indian in blood and colour, but English in taste, in opinions, in morals, and in intellect"), and, on the other hand, movements in England at the same time to enfranchise Jews, dissenters, and Catholics. In other words, both the Anglicization of Indians and the enfranchisement of Jews, dissenters, and Catholics are attempts by the emerging secular state to create deracinated and assimilated "subjects" (eventually "citizens") that are "non-Hindu Hindus," "non-Muslim Muslims," "non-Jewish Jews," "non-Catholic Catholics," and so forth. "Conversion," of course, then becomes a repudiation of all such attempts at assimilation. It becomes a form of dissent, and sectarian groups often become what Immanuel Wallerstein has called "anti-systemic movements." The remaining chapters in the book look at various other instances of conversion, including, in addition to John Henry Newman, already mentioned, the conversions of Annie Besant (from the Church of England to Theosophy), B. R. Ambedkar (from untouchable Hindu to Buddhist), and various other dramatic conversions from Hindu to Christian religion (Ananda Row, Pandita Ramabai, et al.).

The strength of the book is at one and the same time the book's weakness. From the perspective of literary criticism, the book is interesting and informative, providing as it does interesting narrative and biographical accounts from 19th- and 20th-century conversion literature. As a theoretical work about "conversion," however, the book is rather confusing and disjointed, possibly because it is to some extent a pastiche of earlier pieces that were written for other purposes and for other occasions. But even allowing for the obvious problem of organization, it must be said that the book's theoretical orientation is rather limited. While appearing to be sophisticated and au courant in terms of culture theory, poststructuralism, orientalism, and multiculturalism, in fact, the essays for the most part are little more than a series of literary vignettes about Cardinal Newman, Annie Besant, B. R. Ambedkar, and various others. The author's notion of "religion" is uncritical and does not appear to be familiar with important recent theoretical work in religious studies and history of religions. Moreover, the author's notion of "conversion" seldom gets beyond what can only be called an elitist orientation. Attention is given to a few dramatic examples of conversion (Newman, Besant, Ambedkar, Pandita Ramabai, et al.) with almost no attention to what might be called "garden-variety" conversions, or, in other words, the great mass of conversions throughout history, whether in India or elsewhere. As a result, the generalizations about conversion have little theoretical basis beyond the level of eccentric biographical detail and literary vignette. The book, therefore, can hardly be considered an important theoretical treatment of conversion either from the perspective of history of religions or from the perspective of the social sciences.

Also, it must be said that the book is not always well written. The es-

says are sometimes rambling and move in too many directions at the same time. Sentences are frequently Germanic in length, forcing the reader to go over passages a second and even a third time in order to grasp what the author is trying to say. Arguments are often repeated too many times, and tangential asides seem to go nowhere. Many of these problems could have been corrected had the volume been edited more carefully, especially since the essays were not originally written for a single book. In any case, as mentioned earlier, the book is interesting and informative as a collection of narratives and biographical accounts of some fascinating instances of conversion.

*Rendering unto Caesar: The Catholic Church and the State in Latin America.* By Anthony Gill. Chicago: University of Chicago Press, 1998. Pp. xiii+269. $41.00 (cloth); $15.95 (paper).

Paul Christopher Johnson
*University of Missouri, Columbia*

In this tightly argued volume, Anthony Gill applies rational choice theory (RCT) to a fascinating comparative problem: Why did the Catholic Church in some Latin-American states defy authoritarian regimes by expanding ministries to the poor, while in others, such regimes and their associated elites were accommodated? Gill questions the status quo account, wherein the church reformed itself from within, beginning with Vatican II (1962–65). According to this version, reformist popes and prominent bishops in Chile, Brazil, and elsewhere forged a new vision and mission for the church based on a "preferential option for the poor" and an inclusive ecumenicism.

Not so fast, cautions Gill. In fact, such reforms were instituted selectively at best. In the majority of cases, such as that of Argentina, church leadership submitted to strong-arm dictatorships with little protest. How can we explain this wide variation of institutional religious responses to a relatively homogenous political movement across Latin America from 1960–80?

His answer is that in cases where strong third-wave Protestant (Pentecostal) penetrations into the traditional national-religious ideology presented significant competition in the religious "marketplace," the church responded with its own outreaches toward the poor in order to stave off massive defections. In cases without such penetrations, where Catholic monopoly was maintained, the church continued to direct its time, money, and personnel toward wealthy and powerful elites. Hence, in Chile the base community movement (CEB) was given substantial support against the wishes of Pinochet; in Argentina, by contrast, the "dirty war" waged against dissidents was largely ignored by church leadership.

Gill offers, with little problem, an incisive and important corrective to the standard account by offering strong correlational evidence (not caus-

ative) that the church's variable response can be accounted for by viewing it as an interested agent in a competitive religious marketplace. Gill, however, faces a hoarier problem at the end of his account when he attempts to explain the partial return to accommodationist positions in the last decade under the present pope, even in places like Chile and Brazil. Since Pentecostalism has continued to grow rapidly, especially among the economically marginalized, one might expect the RCT model to predict a correlated increase in attention to the poor by the church. But while Pentecostal denominations have continued to make enormous inroads all over Latin America, even the Chilean and Brazilian episcopacies have backed off from Vatican II and Medellín. Why? The easy answer, one Gill refuses, is that John Paul II has sought to consolidate hierarchic orthodoxy by disciplining potential dissidents. Gill proposes that even had this not been the case, church leadership would have returned to the typical pattern of alliance with Latin American national governments. This is because the church, with its weighty bureaucratic baggage, cannot compete over time in the religious marketplace with the comparatively efficient Pentecostal mobilization of resources, and therefore requires state subsidies.

This last argument is less convincing. Limited bureaucratic stability, after all, ought to prove more cost-efficient than the institutionalizing efforts of new movements; this is precisely the sense of Weberian instrumental rationalization arguments. Gill does not engage the problem of how and when rationalization turns into tired bureaucratization and begins to tumble down the cost-efficiency curve.

More seriously, scholars of religion should and will levy some of the standard critiques of RTC applications to religion. The attempt to define what constitutes religious "goods," "preferences," and "interests" seems insufficient. Gill homogenizes these as "compensators" (p. 56), the exchange of present rewards for future or nonverifiable ones. But religious affiliations offer manifold present rewards, and the fact that they are not easily measured or do not fit the model does not mean that they do not exist. Such present rewards include meaning, structure, community, identity, orientation in time and space, and even physical health. Moreover, the attempt to provide a highly generalizable theory leads to other risky homogenizations. At times, Gill slips from axioms about the church into axioms about religion as such (p. 56). Many religions in Latin America, and the Afro-Latin traditions in particular, which Gill lumps generically into the category of "Spiritism," are quite this-worldly in their ritual attention and are structured around quite different "interests" and "preferences" than Christianity, Catholic or Protestant. In the appendix, where Gill explicates RCT for religion, he clarifies by stating that the content of "preferences" is variable, that "it is up to the scholar to assume a set of preferences for actors based upon specific knowledge of the actor" (p. 195). Later, however, and more accurately in my reading, he declares that preferences are specified deductively (p. 228); deductive, one assumes, from the theory with which a religion is approached. The confusion re-

flects the difficulty of applying an economic model of valuation to religious systems.

These are criticisms inevitably provoked by attempts to generate generalizable theories for comparisons of particular historical cases. Gill is to be congratulated for the attempt—audacious, provocative, insightful, and, quite naturally, imperfect as well.

*Seventh-Day Adventism in Crisis: Gender and Sectarian Change in an Emerging Religion.* By Laura L. Vance. Urbana: University of Illinois Press, 1999. Pp. x+261. $39.95 (cloth); $18.95 (paper).

John P. Bartkowski
*Mississippi State University*

In *Seventh-Day Adventism in Crisis,* Laura Vance has produced a monograph that will surely be of interest to scholars who study religion, gender, and social change. Consistent with much current gender scholarship on the emergence of theologically conservative religions, Vance's study reveals how differently history reads when gender becomes a central analytical category for examining religious transformation. This volume aims to address several interrelated questions, including: How did a religious movement in which women initially wielded visionary leadership eventually come to deny women access to many of its most powerful institutional positions? How have large-scale social changes influenced current debates about "women's place" within contemporary Adventism? In fixing her attention on such issues, Vance produces a book that is not simply a historiographical account of shifting gender relations within Adventism—though a focus on that topic alone would have been quite an accomplishment. Rather, recognizing that the best historical research informs contemporary predicaments, Vance combines a backward-glancing eye attuned to Adventism's past with an insightful investigation of present-day gender relations within this religious denomination.

*Seventh-Day Adventism in Crisis* begins by recounting the historical origins of Adventism, a sectarian religion that emerged during the Second Great Awakening of the early 19th century. Special attention is paid to the apparently prophetic visions and writings of Ellen White, an early Adventist thought to have received direct revelation from God, detailing the divine mission of this nascent religious movement. Much of the first half of the book then proceeds to analyze the distinctive—and often paradoxical—facets of Adventist doctrine and practice. For example, Adventists are generally committed to the infallibility of the Bible; yet, at the same time, members of this religious group conceive of divine revelation as progressively unfolding into "present truth." Moreover, Adventism has long decried the excesses of "the world" (e.g., gambling, movie going, and various dietary indulgences) even as it has implored its adherents to affiliate with unbelievers for the purpose of evangelism. The Adventist

challenge of finding one's place "in but not of the world" is very similar to that faced by other theologically conservative religions. Yet, perhaps the greatest Adventist contradiction entails the eventual erosion of women's leadership authority within a religious denomination whose core doctrine was initially defined—or, better, divined—by a female prophet.

In rendering her portrait of Adventism, past and present, Vance avoids homogenizing this diverse and changing religious tradition. Her careful analytical approach reveals how internal cleavages among Adventists themselves emerged historically and continue to surface in light of this religion's conceptualization of an evolving "present truth." Consequently, the first half of Vance's book evenhandedly combines rich idiographic accounts of particular events in Adventist history (e.g., chaps. 1 and 4) with broader analyses of this religion's theological presuppositions and political organization (e.g., chaps. 2 and 3).

Part 2 of this volume focuses on Adventist responses to a series of recent social changes—shifting definitions of gender and sexuality, the recent rise of women's labor force participation, and controversies over women's ordination to the ministry in many Protestant churches. Because Vance has detailed the particularities of this religious subculture so well in the book's first section, she moves deftly through Adventist responses to these various issues—aided, where appropriate, by back references to section one. For example, Vance examines contemporary Adventist support for gender equity in the workplace with an eye on the post-1870 writings by Ellen White, who defended the payment of equitable wages to female employees and became a champion of women's public-sphere participation in Social Gospel movements. Moreover, current Adventist controversies over women's ordination are understood in light of the rich cultural tradition of Adventism. This multilayered tradition contains strands of early Adventist egalitarianism interwoven with more recent accommodations to secularized visions of gender difference. This reading of structural change and ideological diversity within Adventism effectively challenges those who would equate religious conviction—and especially theological conservatism—with an unreflective preservation of the status quo.

Vance has collected and mined a vast array of data to conduct this study. She draws from archival sources, secondary historical treatments, and Adventist pastoral texts. She has also gathered primary data using participant-observation, in-depth interview, and survey techniques. Given the conceptual breadth and methodological triangulation evidenced in this volume, some readers might charge that Vance simply attempts to cover too much ground in one monograph. I do not share that criticism. Although it is easy to envision other works—for example, a more ethnographically focused monograph—that could effectively build on the material in the present volume, this book draws together coherent and compelling narratives from these various data sources. As a result, *Seventh-Day Adventism in Crisis* provides a holistic analysis of a reli-

gious tradition that has undergone great change since its emergence and continues to redefine itself as we enter the next millennium.

*The Gendered Terrain of Disaster: Through Women's Eyes.* Edited by Elaine Enarson and Betty Hearn Morrow. Westport, Conn.: Praeger, 1998. Pp. xi+288. $65.00.

Susan Lovegren Bosworth
*College of William and Mary*

Elaine Enarson and Betty Hearn Morrow have compiled an interesting collection of work on gender and disaster. Their volume establishes "a baseline, reflecting the state of the field" (p. xii) and titillates us to ask about and investigate women in disaster. The chapters collectively suggest that disasters as socially constructed events must be understood within the broader cultural contexts of impacted communities. Disaster researchers have only recently begun to explore what this means to the discipline and to women. The contributions of the volume are interdisciplinary and international, and they reaffirm the need for a gendered look at disaster.

Enarson and Morrow's introduction to the volume locates women in disaster and reveals social structure as a powerful and enduring force even in the midst of physical and social disruption. The introduction is followed by 20 contributions organized into three parts. In the concluding chapter, Enarson and Morrow offer new directions for policy makers, practitioners, and researchers.

Part 1, "Perspectives on Gender and Disaster," summarizes the state of studies of gender and disaster. The three chapters should be read by anyone interested in disaster research. Although I take issue with some of the authors' conclusions, each chapter generates important questions and discussion points. In the first chapter, Alice Fothergill expands the traditional four stages of a disaster event (preparedness, response, recovery, and mitigation) and organizes an overview of the literature in a nine-stage typology (p. 12). Her findings emphasize the importance of gender to understanding social and structural dynamics in each stage of disaster, and the woeful inattention to gender in disaster studies generally. In the second chapter, Robert Bolin, Martina Jackson, and Allison Crist assert that the "functionalist and positivist theoretical orthodoxy" (p. 29) that dominates the U.S. disaster literature has ignored the effects of gender inequality on vulnerability in disasters. As the authors note, this neglect occurs even as feminist theory has been seriously applied to other sub-fields of sociology, and by disaster scholars in the Third World. Joseph Scanlon rounds out the first part of the book with a discussion on myths and why debunking gender myths and recognizing the contemporary social conditions are the first steps in planning appropriate gendered re-

sponses to disaster. I think American scholars are less wedded to a purely functionalist approach than is implied by some of the authors, but the authors' points force us to reflect on the history, current state, and future of disaster research.

The second and third parts of the book are case studies or first-hand accounts of women in disasters. Part 2, "Social Construction of Gendered Vulnerability," emphasizes "how the vulnerability of women is socially and culturally constructed" (p. 53). The contributions represent considerable diversity in content, style, and perspective. For example, one anthropologist writes of her experiences as a survivor in a "living laboratory" (p. 61) after a California firestorm, another uses a comparative case study from Bangladesh to illustrate the effects of domestic structural arrangements on disaster vulnerability and coping capacity. With examples from case studies in Scotland and Wales, two researchers illustrate how the domestic environment of women goes largely unrecognized in the male-dominated and action-oriented "world of emergency planning and disaster management" (p. 93). One chapter looks at the "complex web of factors" associated with postflood disaster stress in three U.S. communities, another at how the daily vulnerabilities of women and children in the Philippines are exacerbated in disasters, and another at domestic violence after disaster. Collectively, these contributions reinforce the theme that the vulnerability of women is "embedded in patriarchal values and their resultant social systems" (p. 54). The policy initiatives included at the end of each chapter reflect issues of social inequality generally and disaster vulnerability specifically.

Part 3, "Case Studies of Women Responding to Disaster," gives a global account of women in disaster. The 11 case studies tap different social and cultural contexts, and they illustrate recent initiatives to address gender issues in disaster. For example, there are descriptions of how relief work in Pakistan was "converted into developmental work" (p. 129) that empowered women and how women organized a coalition in Miami, Florida, to provide a female voice in rebuilding the area after Hurricane Andrew. A formal Red Cross volunteer and a teacher give personal accounts of their disaster experiences. Other chapters look at how women have responded to and coped with disaster. This last section in particular, emphasizes social structure and change, and disasters as catalysts for change.

The book is an important contribution to the disaster literature. As evidenced in the brief description above, it informs practitioners and planners, disaster and social science researchers, and the general reader. The editors provide excellent introductions to each part of the book that will help readers select chapters that are most pertinent to their interests. In the end, the editors succeed in leading us through the "gendered terrain of disaster."

*Gender and Power in the Workplace: Analysing the Impact of Economic Change.* By Harriet Bradley. New York: St. Martin's Press. Pp. ix+250. $55.00.

Sylvia Walby
*University of Leeds*

Gender relations in employment are changing rapidly, inviting important questions as to whether increases in women's employment have entailed significant reductions in gender inequality. Bradley asks whether feminism has improved women's position in the workplace or whether new configurations are just as oppressive as the old. Bradley analyzes five case studies of employment, located within the specific class dynamics of the North-East and the United Kingdom in the early 1900s—an area of high unemployment following the decimation of the old heavy industry of coal and iron and steel, which in turn is located within the dynamics of globalization.

Bradley argues that work is newly important for many women's self identities, that most women can no longer be construed as primarily housewives. She has a subtle account of the interplay of identity and materiality, of the opportunities and imitations for cultural expression provided by the economic restructuring of jobs. Bradley argues that class location should not be equated with class identity, in the manner of some postmodernist writings, arguing that the lack of such an identity does not mean that class relations do not exist. While the author herself is committed to using concepts of class and gender, her interviewees are wary of locating themselves within a class framework, partly lacking confidence as to where they are placed in shifting class hierarchies, partly rejecting the assumptions behind such concepts. Bradley seeks to draw on some of the insights of postmodernism, such as new ways of exploring identity and meaning, while rejecting its flaws. She is highly critical of the importation into feminist analysis of conceptions of power derived from Foucault, which so disperse power that it is not possible to identify specific beneficiaries. She argues resolutely for notions of agency, in order to underpin responsibility. She considers feminism to be about redistribution as well as recognition. The women in her study were not passive victims, whether or not they called themselves feminist.

However, despite increases in women's employment, old inequalities between men and women at work, related especially to occupational segregation and patriarchal attitudes, remain to a considerable, if reduced, extent. Bradley is skeptical of claims that there has been any significant upskilling of women's work, retaining a traditional view of capitalism as hostile to the development of workers' capacities. She endorses the pessimistic version of the flexibility thesis, arguing that work has become more fragmented, more stressful, and more insecure. But while the workers in her sample in the Northeast of England might have such experiences, there is a problem in Bradley's tendency to generalize from a small

and very special sample to the United Kingdom. Bradley does not have a sufficiently representative sample to reject survey based data, which suggests a different picture for the United Kingdom, such as that of Gallie's more optimistic assessment of the upskilling/downskilling question. Similarly, Bradley is skeptical of the impact of education, but a sample where only 6% have university education, in the context where currently 30% of both young women and young men are taking degrees, is not a sufficient basis for generalizing about current developments in Britain.

Bradley is also skeptical of the impact of equal opportunities laws and policies, describing resistance to such policies in the workplace and suggesting that they have more impact on middle-class than working-class women. However, her sample almost entirely excludes part-time workers, despite their making up nearly half of women workers in the United Kingdom and those who are most often found in the least skilled and more manual jobs. Yet, in the mid-1990s, U.K. legal rulings, based on European Union law, have insisted that part-timers should be given the same working conditions, rights, and fringe benefits as full-timers on the grounds that to do otherwise is to indirectly sexually discriminate against a group (part-timers) that is largely female. Such rights, now consolidated in the EU Directive on part-time work, in fact go beyond most U.S. legal practices and affect primarily women in lowly occupations. Working-class women, especially part-timers, thus are benefiting from equal opportunity laws.

The strengths of Bradley's book then lie in its account of the particularities of the intersection of gender and class in specific occupations in the northeast of the United Kingdom and in the subtlety of the interplay between identity and location. She asks the big questions about change, power, and inequality. Her attempt to frame these case studies within a wider context is both a strength, in that it situates them more broadly, and a weakness, in that she does not really have a sample sufficiently large or representative to justify the more general of her claims. It is hard to both use case studies and to address the wider questions of changes in capitalism, gender relations, and globalization.

*The Two-Headed Household: Gender and Rural Development in the Ecuadorean Andes.* By Sarah Hamilton. Pittsburgh: Pittsburgh University Press, 1998. Pp. v+296. $50.00 (cloth); $22.95 (paper).

Amy Lind
*Arizona State University*

Do women and men participate, in egalitarian terms, in the management and distribution of household resources in the Ecuadorean Andes? Are gender relations in Andean households hierarchical and unequal, as in western societies, or complementary, as some researchers have argued? Do women share decision-making power with men in their households

and, by extension, in their communities? These are some of the questions Sarah Hamilton attempts to answer in her ethnography of household relations in one Ecuadorean village.

The author utilizes both quantitative and qualitative research methods to examine household and community relations in Chanchaló, "an agricultural community of some 135 households located [in] . . . central Ecuador" (p. 37). In chapter 1, "Gender and Rural Development," she provides an extensive review of the literature on gender and development studies and rural Andean studies, drawing in particular from debates in anthropology. The review is useful but not as interdisciplinary nor as contemporary as I would like. In particular, it lacks a more engaged analysis of recent postmodernist critiques of development, Andean studies, and feminist research methods. These areas of research provide important critiques of objectivity, which would strengthen both the literature review and the author's ethnography.

In chapter 2, "An Introduction to Chanchaló," the author draws from secondary sources and from data collected for a National Science Foundation (NSF) study to situate Chanchaló in the context of regional and national development. In chapter 3, "Managers, Mothers, Maiden, and Matriarch: Five Women and Their Families," the author introduces five of the women included in her larger ethnographic sample of 10 women. In this chapter, she describes the daily lives of these five women, focusing on their reproductive and productive activities in the household, in agricultural development, and in the community. In chapters 4–7, the author combines data from the quantitative and qualitative studies to examine various aspects of the Andean two-headed household: women's productive and reproductive work, women's control of household economic resources, structural and ideological foundations of the two-headed household, and gendered labor markets.

It contrast to other studies of gender and development in the Andes, Hamilton argues that women's egalitarian status in the household and in Chanchaló have not been undermined as a result of "planned" and "unplanned" processes of development. Rather, the author asserts that ". . . [i]n Chanchaló . . . women and men share control of their households' most important economic resources in an egalitarian manner," despite the historical integration of Chanchaló into capitalist markets (p. 182). Furthermore, she argues that "the two-headed household is a product of both tradition and change," rather than merely a product of tradition, as might be argued from a romanticized perspective (p. 209).

One strength of this book is Hamilton's ethnographic account of 10 women in the Chanchaló community. Another strength is the author's ability to integrate her own microlevel narratives with broader policy analysis. The author's questions about gender relations in the household are not new, as feminist sociologists and economists have theorized extensively about gender relations within households, and researchers have long argued about the usefulness of notions of gender complementarity versus gender (in)equality in describing Andean social relations. Despite

this lack of originality, Hamilton's original contribution lies in her ability to document in a complex fashion the nuances, contradictions, joys, and difficulties of men's and women's lives in Chanchaló households.

The author provides a critique of the economistic aspects of conventional gender and development scholarship, although she does not go far enough in rethinking fundamental concepts such as the economy, the household, and the family in her own examination of Chanchaló households. Her project may have been enriched by scrutinizing more carefully some of these concepts central to the literature and to her own way of conceptualizing Andean gender relations—and by broadening the research question to include other social and cultural factors such as non-economic relations of power within households (an area which she discusses primarily in chapter 6, "The Power of Balance: Structural and Ideological Foundations of the Two-Headed Household."

Despite its flaws, *The Two-Headed Household* is a rich ethnographic account of household relations. It will benefit researchers and students interested in learning about gender, household relations, and economic development in cross-cultural perspective.

*From Duty to Desire: Remaking Families in a Spanish Village.* By Jane Fishburne Collier. Princeton, N.J.: Princeton University Press, 1998. Pp. xi+264. $55.00 (cloth); $17.95 (paper).

Arlene Skolnick
*University of California, Berkeley*

In the 1980s, Jane Collier returned to "Los Olivos" (not its real name), the small Andalusian village she and others had studied 20 years earlier. Dramatic economic, social, and political changes had taken place between the two visits. Industrialization and the development of commercial farming had transformed the village and its social institutions. Agriculture was no longer the basis of the local economy. Occupational achievement had replaced land ownership as the source of status. With the end of the Franco dictatorship and its replacement by a democratic government, Spain itself had been reconfigured as a nation and as a culture. Collier examines the social and cultural changes that had remade Los Olivos, focusing on the family and on the development of a "modern subjectivity." The courtship and marriage system that Collier had observed in the 60s had disappeared. Extended courtships that spanned years no longer prevailed. Also absent was the series of rituals that had marked their progression. ("Those earlier customs were not normal!" declared one woman to Collier.) Young people now married for love, with parental role in the choice of a spouse comparably diminished. The guiding principle of marriage was now companionship and intimacy. Partnership had replaced patriarchy. Contemporary parents valued affection

more than obedience from their children. All this had happened as standards of living increased dramatically.

As Collier's subtitle reminds us, the core of her book is a richly detailed ethnography of social change. We learn how people spent their days and evenings in the 1960s and 1980s, how they dressed, cared for babies, did the wash, what they said about their lives and the lives of others, what they complained about, and the dilemmas they faced at the two periods. In the 1980s, people told Collier what they thought about the changes they had witnessed in their lives.

But there is another more theoretically ambitious book here, the one indicated in the main title—*From Duty to Desire*. This volume relates the author's struggle to make interpretive sense of the ethnographic material. Had this book been written between the 1950s and the 1970s, the concept of modernization, or one of its theoretical variants, would have supplied the framework of interpretation. The families in Los Olivos would have been said by Talcott Parsons to have undergone "structural differentiation?" or to have exemplified what William J. Goode called the "world revolution in family patterns" (New York: Free Press, 1973), in his book of the same name; or as historians have written on "traditional" families based on obligation and patriarchal authority that had been transformed into "modern" families based on domestic affection. But this is a book of the 1990s, and our understanding of modernity has become more complicated. The "modernization" narrative has become unfashionable, even suspect.

Thus, Collier says she did not want to write "the standard modernization story of progress from constraint to freedom" (p. 10). Instead, she focuses on the contrasts between two "mentalities"—"traditional" people tell stories about rules and obligations that seem to them to make their lifestyles feel "handed down," while moderns tell stories about "thinking for themselves" and making choices. Although she says she wants to avoid accusing moderns of false consciousness, Collier believes that they are deluding themselves.

She interprets her findings from a Foucaultian perspective. Thus, she concludes that "people from Los Olivos may have given up the visible chains of 'traditional society' for the constant, if less visible, surveillance by 'modern disciplinary apparatuses'" (p. 10). She does not believe that the villagers had "significantly more choices open to them" in the 1980s than they did in the 1960s, "even though they talked and acted as if they did" (p. 26).

Collier's postmodernist stance raises serious questions about the theoretical and methodological, as well as political implications of the Foucaultian vision of modern society, especially when applied to specific historical circumstances. Certainly it would be naive to assert that we moderns are totally free of social control and coercion. But as applied to particular people living in a specific historic time and place, Foucault and other postmodern or poststructuralist theorists risk seeming unrealis-

tic, even naive. After all, when Collier's initial research was completed, Spain was governed by a Fascist dictator who affirmed tradition in the interests of social control.

Collier quotes from a 19th-century novel *Doña Perfecta*, to open each chapter and also to illustrate differences between the "mentalities" of tradition and modernism. In a tragic tale somewhat along the lines of Romeo and Julier, young love clashes with hypocrisy, greed, and blind adherence to ancient customs. The author, Galdos, clearly prefers the side of modernity and the young lovers. Collier dismisses him as a "bourgeois liberal." Thus, she combines a relativistic and, at times, appreciative stance toward the world of Doña Perfecta, with a highly critical stance toward liberal modernity, the epistemological inconsistency Calhoun has called a "widespread anthropological neurosis" (*Critical Social Theory* [Blackwell, 1995]).

Like Foucault, Collier provides no clear guidelines for distinguishing between degrees of social control and coercion or between legitimate and illegitimate power. It is scarcely imaginable that Collier really wants to conclude that people in post-Franco Spain are no more free than they were under the dictatorship or that fascism and democracy are simply different but morally comparable systems, but that is where her analysis seems to lead.

Despite my differences with Collier, I strongly recommend this book. It is a beautifully written narrative of social and cultural transformation that would be useful in a number of courses. Further, the author's theoretical arguments with herself and others make it good to teach with also.

*Fathers under Fire: The Revolution in Child Support Enforcement.* Edited by Irwin Garfinkel, Sara S. McLanahan, Daniel Meyer, and Judith Seltzer. New York: Russell Sage Foundation, 1998. Pp. xiii+351.

Frances Goldscheider
*Brown University*

This important and highly informative collection of studies on nonresident fathers and child support should be of great value to scholars and policy makers, and perhaps in advanced college courses. In this era of reduced state expenditures and large support awards that accrue even when men are unemployed or in jail, it raises a series of pertinent questions, which frame the collection's three sections. The first asks what the policies are and who the fathers are. The second presents five rich research studies asking how child support enforcement affects fathers. The final section asks whether and what we should do to help fathers.

The answers are often counterintuitive, reinforcing the importance of research for public policy. For example, men do not appear to reduce their incomes to avoid paying child support (chap. 4). Regarding de-

mographic outcomes, enforcement does reduce repeat out-of-wedlock childbearing (chap. 7), although not childbearing in a new marriage. However, marriage is also deterred, so the authors of chapter 5 estimate that a 10% increase in collections is a wash for existing children's well-being, because gains to biological children are lost in the decline in remarriages involving stepchildren.

The studies regularly distinguish between poor and nonpoor fathers, which is particularly helpful. For better-off fathers, the results suggest that paying more will not impoverish them and any new family members substantially. Certainly, their economic well-being would remain above that of custodial parents, thus reinforcing the image of the "deadbeat dad." We also learn, however, that between 30% and 40% of all nonresident fathers earned less than $6,500 in 1995 dollars (p. 53). Given that the major targets of child support enforcement are fathers of children on welfare, this research shows that the emerging patchwork of child support enforcement lacks fairness and legitimacy. In the words of the editors in the concluding chapter: "To insist that these men pay as much child support as a man with a full-time, full-year minimum-wage job is unduly onerous. . . . The inevitable result . . . is the accumulation of child support arrearages, periodic jailing, and the buildup of hostility and resentment towards mothers and children as well as government authority" (pp. 332–33). The hostility is amply documented in chapter 8, which is on increasing fathers' access to their children.

A perhaps equally important contribution of the volume is the examination of the ways standard surveys systematically underrepresent nonresident fathers (chap. 2). They estimate that at least 36% of the "missing fathers" gap does not result from underreporting. Rather, it is the result of the nonrepresentation of persons in group quarters (primarily those in prison and the military) and the underrepresentation of men who are poorly attached to a permanent residence (the bulk of census undercount). The gap might even be less. Many fathers who are reported to have no contact might be dead. We need to apply survival rates such as those generated for inner-city men by Geronimus to see how large this group might be. Another reason the numbers might not match is that more men may have fathered children with many women rather than vice versa, as the evolutionary biologists hypothesize. The authors considered this but ignored the possibility that absent parents report only some of their children. I look forward to the day when the care we routinely apply to surveying women's contraceptive use, abortion, and breastfeeding duration is applied to the challenges of men's fertility histories.

The final substantive chapter asks a series of moral questions about child support, such as should poor parents pay a smaller proportion of their income in child support than nonpoor parents? And should a nonresident parent get a reduced child support award after having additional children, thus reducing the well-being of earlier children (much as intact families routinely do to their first children when they continue)? But perhaps we should be asking even more searching moral questions, such as

how can the country that styles itself as the richest in the world invest so little in its children and harass their poor parents so much?

The research reported in the volume, of course, raises more questions than it answers. Many more questions, as I have suggested, should also be raised as we confront the detritus of the decline in male-female commitment in the context of the final frontier of the gender revolution—the role of men in families. I welcome this book's focus on "fathers under fire," particularly from a class perspective. However, from a gender perspective, it reinforces sexist images of men as fathers. It ignores both the full and partial-custodial versions of fatherhood and is obsessed with violence. Perhaps the most comforting finding in this period of rapid and chaotic change is that (above a certain level) increased effectiveness of state programs appears to reduce parental conflict.

*Marriage in Men's Lives.* By Steven L. Nock. Oxford: Oxford University Press, 1998. Pp. viii+165.

Linda J. Waite
*University of Chicago*

*Marriage in Men's Lives* is a courageous and innovative book: courageous because it tackles a politically and socially charged issue—marriage as a social institution—in a time when texts on the family portray marriage as just one of any number of equally valuable lifestyle choices (see *Close Hearts, Closed Minds: The Textbook Story of Marriage* [1997] by Norval Glenn); innovative because it looks closely at the ways in which a key social institution affects individuals, in this case, the way that marriage affects men.

Even as sex differences within marriage have diminished, the role of husband still plays a unique function in the lives of men. Steven Nock argues that adolescent boys face challenges in becoming men that adolescent girls do not face in becoming women. According to Nock (pp. 6–7), "Masculinity is precarious and must be sustained in adulthood. Normative marriage does this. A man develops, sustains, and displays his masculine identity in his marriage. The adult roles that men occupy as husbands are core aspects of their masculinity." The behaviors expected of married men as *husbands*, according to Nock, are the same behaviors expected of husbands as *men*. So getting married and successfully doing the things that husbands do allows men to achieve and sustain their masculinity. Nock defines what he calls "normative marriage," distinguished by free choice, maturity, heterosexuality, husbands as head of the household, fidelity/monogamy, and parenthood. His measures of these theoretically important dimensions are clever *and* thoughtful, an unusual combination.

Nock argues that if marriage provides a mechanism through which men establish and maintain their masculinity, marriage should have con-

sistent and predictable consequences. He reasons that normative marriage will have different consequences than other forms of marriage. He looks for evidence of the impact of marriage on men's adult achievement, their involvement, participation and engagement in social life and organizations, and their expressions of generosity or philanthropy.

Nock argues that marriage *causes* men to become more successful, to participate in social life, and to become more philanthropic. This is, in today's climate of extreme caution about causal relationships, a bold claim. He tests it using data from the National Longitudinal Survey of Youth and fixed effects models, to separate changes that accompany aging from those that happen uniquely at marriage. To measure achievement, Nock uses annual income, annual weeks worked, and occupational prestige. He measures social participation with time spent on housework, social contacts, and organizational involvement; and he measures generosity with gifts to nonrelatives and loans to relatives and nonrelatives.

To summarize his results too briefly, when men marry, their achievements rise on all measures; they reduce their time in housework; increase their contact with relatives, church services and church events, and co-workers; and decrease contact with friends and time in bars. When men marry, they give fewer and smaller gifts and loans to nonrelatives and more and larger loans to relatives. Nock also looks at changes in each of the measures of adult achievement, social participation, and generosity with changes in each of the dimensions of normative marriage. He finds, generally, that moves toward normative marriage increase achievements, social participation with family and religious organizations, and generosity to relatives. Changes toward more normative marriage also reduce men's time in housework, their social contacts with friends, and social events in bars.

Nock concedes that coercion and gender inequality underlie traditional marriage but argues that "dependency" rather than inequality might be a better description. Mutual dependency is a powerful cohesive force holding marriages together and increasing commitment. In fact, marriage gets much of its power from the interdependency that allows each spouse to specialize. Can perceptions of inequality, typically *women's* perceptions of inequity, be reduced without weakening or destroying the solidarity of marriage? And can marriage still serve to establish and reinforce men's masculinity, which they do by being protectors and providers for their families, when men share the role of family provider with their wives? Nock thinks so.

Nock advocates a new, equitable model of marriage—a new normative marriage. This new marriage bargain would call on men to work, but not require that they make more than their wives. As part of this bargain, husband and wife must recognize themselves *both* as dependents in their joint project—the family—even if they have very different incomes or one has no income at all for a period. But to achieve this alternative vision of the "new family," dependencies must be freely chosen, not co-

erced. Nock (p. 136) points out that this new arrangement requires nothing more than "a way of seeing things differently."

This book sets new terms for the debate about the future of the family, especially about what it means for men and where they are without it. I recommend it highly.

*Gender and American Social Science: The Formative Years.* Edited by Helene Silverberg. Princeton, N.J.: Princeton University Press, 1998. Pp. x+334. $55.00 (cloth); $18.95 (paper).

Patricia Madoo Lengermann
*George Washington University*

This collection of 10 essays adds to the burgeoning literature about the role played by a politics of gender in patterning American social science. The formative years of American social science, 1870–1930, coincided with a large-scale mobilization by privileged-class women for greater empowerment, which privileged-class men met with accommodation and resistance. The emergent social sciences became an important site for this struggle.

This volume of informed, detailed, and accessible essays allows us to trace the effects of this politics of gender in several social science disciplines: economics, anthropology, home economics, political science, and sociology. The authors choose different strategies for describing this gender politics. Several do so by recovering the inspired agency of individual women who worked mightily to construct a social science, only to be written out of histories of the social sciences by a later, regrouped, male "power class." Kathryn Kish Sklar charts Florence Kelley's commitment to a critical research agenda in the production of *Hull-House Maps and Papers;* Dorothy Ross discusses June Addams as a major interpretive theorist; Delsey Deacon describes Elsie Clews Parsons's radical challenge to American norms about family, sexuality, and mannered society; and Guy Alchon traces the blend of scientific faith and reformist passion in Mary Van Kleeck's foundational work on the sociology of work and organizations. In a more generalized approach to women's agency, Kamala Visweswaran rejects the commonplace idea of early women anthropologists building careers through the patronage of sympathetic and influential men, arguing instead for the independent and feminist-supported careers of early women anthropologists. Portraying Cornell's School of Home Economics, Nancy Berlage presents a generalized account of women's ambition, discipline, institution building, and research achievement. In contrast with this focus on women's agency, Nancy Folbré describes the resistance of male economists who discriminated against women in graduate work and professional employment, remained loyal to male-dominated unionism, held to arguments about a "family wage" for men and the ephemeral and subversive nature of women's wage work, and consis-

tently discounted the economic significance of research done by women—
like Edith Abbott and Sophonisba Breckinridge—on women's paid and
unpaid work. The two chapters on political science use key theoretical
texts of the period (Mary Deitz and James Farr) and Progressive munici-
pal reform (Helene Silverberg) to show how masculinist concepts of the
state, politics, and citizenship were established in defiance of liberal femi-
nist claims for women's equal political rights and of cultural feminist
claims for a "municipal housekeeping" agenda.

The essays taken together reveal several common themes: the social
settlement as a crucial site in women's social science (as the university
department became for male social science); the framing context of femi-
nism, Progressivism, the women's club movement, and Social Gospel re-
form agencies like the Unitarian church and the YWCA; the consequent
commitment in women's social science to an agenda of ameliorative so-
cial change (contrasting with the male emphasis on a value-neutral, non-
interventionist social science); the women's practical inventiveness in em-
pirical research, which they saw as a tool for uncovering social injustice
and for persuading publics of the necessity for reform policies; and the
women's ingenuity in creating spaces in which they could earn a living,
practice their version of social science, build a public reputation, pursue
a career, and have an impact on public policy—spaces like the settle-
ments, the public lecture circuit, the Consumers' League, the YWCA, the
Russell Sage Foundation, the Chicago School of Civics and Philanthropy,
the Federal Children's Bureau, the Department of Labor's Women's Bu-
reau, and, within the academy, programs like household administration,
home economics, social administration, social work, and, to a degree,
management science.

Perhaps what is most striking and generalizable about these essays is
what Ross felicitously calls "the historical contingencies of disciplinary
boundaries and definitions" (p. 236). We today stand so surely and so
insularly in our various social science fields—fields whose boundaries
emerged out of the academic politics of this earlier time—that it is hard
to imagine these earlier women working without the framing structure
of these disciplinary institutions. Yet that this is so is shown in how the
authors in the collection claim the same women over and over again as
"founders" of their various fields.

Yet, *Gender and American Social Science*, though stimulating to read,
is troubling in several ways. One is the editor's claim that all this is new
knowledge and therefore news to the informed reader—all the general-
izations made above could have been made by any sociologist working
in and paying attention to trends in scholarship on the history of sociology
without reading this collection. Another is that while all the other disci-
plinary studies are done by practitioners in the discipline, the sociology
essays (with the exception of Deacon on Parsons) are done by historians—
leading to a lack of the nuance an "insider" perspective might have pro-
duced. Third, the collection needed to adhere more regularly to the edi-
tor's proclamation that "this volume . . . is not a book about women in

the social sciences . . . [but about] the many ways in which gender rela-
tions . . . shaped the development of the American social sciences" (p. 3).
While this goal is achieved in some essays, notably those on political sci-
ence and anthropology, *all* of the essays about sociology are about specific
women's contributions to the developing field of thought.

*Sexology in Culture: Labelling Bodies and Desires.* Edited by Lucy Bland
and Laura Doan. Chicago: University of Chicago Press, 1998. Pp. ix+236.

*Sexology Uncensored: The Documents of Sexual Science.* Edited by Lucy
Bland and Laura Doan. Chicago: University of Chicago Press, 1998. Pp.
xv+261.

John H. Gagnon
*State University of New York at Stony Brook*

These two volumes are a valuable attempt to join together a set of late
20th-century essays on the history of the field of sexology (*Sexology in
Culture*) and the late 19th- and early 20th-century sexological literature
to which the essays refer (*Sexology Uncensored*). The first volume is com-
posed of 13 academic essays, which are grouped into four sections that
focus in turn on the dangerous effects of sexological ideas, how sexology
linked bodies and identities together in new configurations, the sexologi-
cal attachment of new desires to old practices, and the ways in which
culture created perversions. These section divisions seem somewhat post
hoc and do not precisely map on the eight more conventionally labeled
sections (e.g., gender and sexual difference, reproductive control, homo-
sexualities) into which the volume of documents is divided, though the
documents selected are nearly all intimately connected to one or another
of the essays. The volume of documents contains 64 quite short selections
that were originally published from the mid-1880s to the late 1930s. The
selections are about equally distributed by decade from 1890–1939 and
represent the work of 39 "authors" (though there are fewer sources since
a considerable number of selections are from the same work). We are
told by the editors that a major criteria for the selection of the documents
was their inaccessibility, hence the use of "uncensored" in the title. Based
on this criteria of general availability, the selections do not include any
works from Freud (though the editors point out his intellectual connec-
tions with the sexological movement). This claim of "making available
the unavailable" poses a puzzle for this reader: Did this specific collection
of documents determine which essays were written or did a set of pre-
existing essays largely determine which documents were selected? Both
processes appear to have been operative, but it is the latter that seems
to predominate. From my reading, a set of contemporary essays were
written based on a much larger body of sexological literature, which in
turn became the basis for a set of exemplary selections.

As a body of work, these essays and documents are a useful empirical response to Foucault's challenge to the received view that sexology played an important role in liberating Euro-Americans from what was thought to be the sexual straitjacket of Victorian prudery. Foucault argued that sexological "discourse" was part of an outpouring of sexual writing and talking that he reports (with limited evidence) characterized the latter two-thirds of the 19th century. Rather than liberating people from prudery, Foucault asserts, sexologists were simply one battalion of the new style army of certified experts in the disciplines who were creating a socially and mentally "classified" humanity. For the most part, these essays represent an attempt to get past both the simplifications of the "liberationist" and Foucaultian visions of sexology (though Foucault is treated quite respectfully) to a more detailed and complex reading of sexological texts. In these essays and the documents that accompany them, sexology turns out to be a contradictory and multipurposed intellectual movement, sometimes advancing ideological arguments that are popular among the essayists, other times taking for granted 19th-century ideas about race, eugenics, and women.

The focus in both the documents and the essays is British. The sexological works themselves are mostly from English authors (12 of the 64 are by Havelock Ellis, two others are responses to his work) or are from works translated into English (mostly from the German—primarily Otto Weininger and Magnus Hirschfeld) that were influential in England. The essays that examine these sexological works are by Anglo-American authors and deal with the impacts of the sexological writings in the British context. This leads to a certain insular quality to both the selections and the essays. While this insularity tends to obscure the pan-European nature of the sexological enterprise, both as research and reform, it appears to offer a certain cultural continuity since contemporary English speakers are writing about "their" cultural past. I write "appears" since this implicit sense of continuity may be somewhat of an illusion.

The literary critic George Steiner has written extensively about the problem of understanding works that are distant from us in time, in language, or in culture (with a small and capital "C"). In this selection of documents, those written prior to World War I have a particularly alien quality about them, as if they embodied understandings and classifications that have slipped over the historical and cultural horizon. One wonders what an American college sophomore who believes that John Kennedy is the name of his or her junior high school will make of Kraft Ebing's strange case histories written in turbid prose. The more recent works of the 1920s and 1930s have a more familiar sound, raising issues that have a resonance in recent history. As even the past grows more distant to understanding, it is clear that what is enticing to contemporary scholars about the sexological past is not what the 19th- and early 20th-century sexologists thought was enticing about sexuality. The selections from the past and the essays are ordered into a late 20th-century set of interests (e.g., race, eugenics, gender or "bodies," "difference," "subjects"),

questions that the sexologists thought about in quite different ways or did not think about at all (here is the importance of the "absence" to postmodernists or why our ancestors did not write about topics that consume the narcissistic present).

This absence of relation between the now and the then and the hegemonic ways in which the now makes use of the then is what most impresses me about these two volumes. The essays themselves are nearly all thoughtful and carefully researched, but they are very contemporaneous, in some cases postmodernist in their interpretative impulses. They often do not attempt to make the documents of the past understandable in their own terms. Why would sexologists write in the way they did? Why would they generate these strange categories? Who were the people who came forward and talked to them? Why is the case history the storytelling document of choice? The cutting up of the past into text bites works against the essayists' interests in capturing the wholeness of the social and cultural situation that produced this kind of thought.

Even with these strictures, these volumes do substantially advance our understanding of the complexities and contradictions of the sexological tradition. And they give it a place in the history and the sociology of science.

*Disciplining Reproduction: Modernity, American Life Sciences, and "the Problems of Sex."* By Adele E. Clarke. Berkeley: University of California Press, 1998. Pp. xvii+421. $45.00.

Steven Epstein
*University of California, San Diego*

This is a story, as Adele Clarke tells us at the outset, about women and men, about rats and guinea pigs, about medical schools and agriculture research centers, about the politics of birth control and the funding practices of private foundations, about stallion urine and vaginal cells. More to the point, it is the story of how a wildly heterogeneous assortment of actors, institutions, and laboratory practices came together in the historical formation of a new scientific field in the United States: the modern reproductive sciences. Between 1910 and 1963, reproduction was "disciplined," Clarke argues, and the Foucauldian resonances of the phrase are fully intended. The consolidation of a new scientific discipline, cobbled together from varied sources but forming its own knowledge base, journals, and professional associations, proceeded hand in hand with diverse efforts to harness what Foucault called "biopower"—the ability to control bodies and populations through the scientific management of sex.

As Clarke explains, the reproductive sciences were "late" in developing, and they have continually struggled to evade the taint of illegitimacy— the perception that there was something not quite kosher about the scientific scrutiny of sexual parts and practices, mixed with the fear that the

science of reproduction might usher in a dystopic "brave new world." But scholars in science studies have also, for whatever reasons, been slow to take up the reproductive sciences as a topic of scholarly investigation. Clarke's is therefore the first comprehensive sociological and historical account of the rise and maturation of the reproductive sciences; and, in it, the full scholarly apparatus of recent social and cultural studies of science, technology, and medicine is displayed and deployed. This is a book that rests on exhaustive scholarship, and, at least to the casual reader, few stones appear to be unturned. At the same time, *Disciplining Reproduction* is chock full of terms and concepts from science studies, and Clarke is adept at demonstrating their applicability in explaining the vicissitudes of the development of an important scientific field.

In many ways, the story Clarke tells is about social and scientific boundaries: how they are erected, traversed, torn down, and built up in new ways. On the one hand, the reproductive sciences were formed when a pre-existing unity—"a densely interwoven nexus of problems in the life sciences grouped around 'evolution/heredity/development/reproduction'" (p. 67)—was yanked apart to form three different scientific domains: genetics, developmental embryology, and the reproductive sciences. On the other hand, the reproductive sciences were able to take shape only through the marshalling of intellectual and practical resources drawn from three other, relatively autonomous fields: biology, medicine, and agriculture. In describing this "cross-professional intersectional scientific enterprise" (p. 130), Clarke tracks the simultaneous unfolding of events within biology, medicine, and agriculture, and shows how reproductive scientists emerging from those fields gradually forged a new professional identity.

But this is not just a story of scientists and their work, for Clarke is at pains to emphasize the constitutive role of funders, and of diverse sorts of birth control advocates, in the very establishment and trajectory of the scientific field. Linked to an "illegitimate" and "controversial" domain of scientific inquiry, reproductive scientists were, until the mid-1960s, peculiarly dependent on their funding sources for legitimacy. But in the end, scientists proved adept at "capturing the sponsor" (p. 90)—convincing funders to support the basic zoological and endocrinological research that reproductive scientists wanted to pursue, thereby marginalizing research concerned with the organization and expression of human sexuality. Somewhat similarly, reproductive scientists gradually succeeded in working out a "quid pro quo" with birth control advocates. In an arrangement that in fact awarded to the scientists the "definitional authority" over contraception, reproductive scientists agreed to devote their energies to developing birth control technologies, but only on the condition that they be allowed to emphasize basic research on "modern," technologically advanced forms of contraception, to the exclusion of scientific attention to more simple, chemical and mechanical means of preventing pregnancy. In her consideration of the complex negotiations between scientists, funding sources, and birth control advocates, Clarke consistently

underscores two important observations. First, history is contingent, and the development of the reproductive sciences could well have taken other paths—and the potential differences matter, both scientifically and politically. Second, the boundaries between science and other arenas of social life are sometimes exceptionally fluid, and fixing those boundaries—demarcating a domain recognizable as "science"—may as often be the object as the prerequisite of scientific labor.

Although Clarke borrows widely from different theoretical positions within social and cultural studies of science, technology, and medicine (including feminist studies of science), her purpose here is to argue for the utility of one approach: the "social worlds" school, derived from symbolic interactionism and, more immediately, from the work of Clarke's mentor, Anselm Strauss. This approach focuses scholarly attention on specific "arenas of action" that are composed of multiple social worlds—more or less distinct "communities of practice and discourse." It is an approach that seems ideally suited to the case study at hand, for it directs the analyst to trace the activities of all the various actors whose work contributes to the production of knowledge and meaning, and to emphasize the complex negotiations that must be conducted between "worlds" in order for knowledge and practices to become institutionalized and legitimated.

This book therefore functions well on multiple levels—as a thick empirical account of disciplinary emergence; as historical sociology of the management of life, sex, and reproduction in the early- and mid-20th-century United States; as a proof-of-the-pudding proffer in support of the viability of the social worlds approach within science studies; and as a vehicle for rumination on a whole host of big themes like modernity, legitimacy, and the politics of gender and sexuality. And yes, I think it may try to do too much. At times, I felt overwhelmed by Clarke's apparent desire to list, enumerate, and catalog every last fact. At times, I thought Clarke might have been better off addressing fewer themes more systematically. Nevertheless, the book fills an important gap in our understanding of the modern life sciences. It reflects meticulous research and capacious intellectual curiosity—qualities that should bring *Disciplining Reproduction* to the attention of many readers.

*Prostitution, Power and Freedom.* By Julia O'Connell Davidson. Ann Arbor: University of Michigan Press, 1998. Pp. 232. $49.50 (cloth); $19.95 (paper).

Lynne Segal
*Middlesex University*

"I am invisible," the prostitute "Desirée" told researcher Julia O'Connell Davidson. On this occasion, it was not the "punters" (her clients paying for sex) before whom she felt invisible, but the feminist academics and

activists she tried to address at a conference in London a few years ago on the theme of prostitution and violence.

Davidson's empirically wide-ranging account of sex tourism is not any "postmodern" attempt to produce new and subversive reflections in which "sex workers" (like the North American "Scarlot Harlot" or "Annie Sprinkle") see themselves as transgressive "sexperts," subverting traditional sexist significations and transforming social institutions. Nor does Davidson locate her research within the currently fashionable "third way" perspective, seeking some middle ground within the controversies between "sex radicals" and "radical feminists," which have loomed so large in feminist discussion of prostitution over the last few decades. She focuses unblinkingly upon the devastating cruelties, marginalization, and violence suffered by the majority of workers in the prostitution trade, whether sold into bondage in brothels in India or Bangkok or working the streets in Cape Town, the Dominican Republic, or elsewhere.

O'Connell Davidson researches the lifespans and collects narratives on the daily experiences of prostitutes, their "pimps" and clients, alongside those of brothel keepers and others involved in the sex trade. From her findings, it is clear that while some sex workers may indeed find personal as well as financial satisfaction in their work, the majority of people selling their bodies for sex do so in desperate social conditions that involve multiple oppressions, offering them little, if any, choice. Because prostitution is not seen as a positive choice by the mass of people who practice it, O'Connell Davidson rejects one recent conceptualization of it as a sexual "orientation" or preference, needing to be incorporated into a pluralist sexual ethics of mutual respect.

Paying close attention to the economic, political, and interpersonal inequalities underlying the engagements of women, children, and some men and boys in the prostitution trade, O'Connell Davidson illustrates the contrasting, but rarely absent, levels of constraint experienced by different groups of prostitutes. Although well aware of the tight bonds between the abuse and mistreatment of prostitutes and surrounding cultures that control, belittle, and stigmatize women's bodies and sexuality, O'Connell Davidson sees prostitution as more than a general effect of male domination. Her analysis can help to clarify some of the conceptual confusions that divisions within feminist scholarship have generated, whether debating the nature of "heterosexuality" and "violence" or discussing issues of "consent," "sexual freedom," and freedom from sexism.

In analyzing the limitations of liberal "contract" theory, for example, O'Connell Davidson is respectful yet critical of the groundbreaking work of Carole Pateman, where she positions "the sex act" as definitive of women's subordinate contractual position, rendering prostitution as no more than an example of the everyday operation of gendered power relations. She is more dismissive of the rhetorical provocation of MacKinnon and Dworkin, reducing all heterosexual engagement to an instance of women's bodily colonization, or violence against women—with the penis itself

here identified as a weapon. Against the totalizing tendencies in such feminist analysis, O'Connell Davidson argues that what we need is much more rather than less focus on the multifaceted *specificity* of prostitutes' subordination, including its "racialized" and ethnic dimension. And this is precisely what her own observations and in-depth interviews offer up to her readers.

O'Connell Davidson tries to steer a steady path through the differing political debates on prostitution, her eye at all times on the most useful strategies for collective political action to reduce the exploitation, abuse, and violence currently endemic to the institution of prostitution and for clarifying the arguments that might force governments and global financial institutions to acknowledge their own responsibility for the ways in which welfare restraint and enforced policies of structural readjustment fuel the worst indignities of prostitution. This book succeeds in its goal of providing a comprehensive survey of a generation of research and debate on prostitution, with many pointers for political action or "realistic reform." It can be used as a guide for activists and policy makers attempting to highlight the bleakest aspects of prostitution, while pointing to the immense obstacles in the way of even the smallest reforms. Some, like me, may feel much less dismissive than O'Connell Davidson of the potentially subversive effects that freely chosen or self-conscious "prostitute performance" may enact in challenging the sexist and very frequently racist abjection of the figure of the prostitute. However, in a world that is becoming more unequal by the day, it is indeed hard not to focus, as this book does, on the constraints and powerlessness that structure the lives of most prostitutes worldwide.

*Making Sense of Prostitution.* By Joanna Phoenix. New York: St. Martin's Press, 1999. Pp. xi+200. $65.00.

Julia O'Connell Davidson
*University of Leicester*

Feminists have long been divided on the question of whether prostitution should be understood first and foremost as an institution that reflects and consolidates men's sexual and political power over women or as a form of rational economic action on the prostitute's part. Caught in this either/or framework, many who write on the subject feel the need to stake out a position and defend it against incursions from those in the opposing camp. *Making Sense of Prostitution* represents an attempt to move beyond this sterile dualism by examining the meanings that women attach to the contradictory effects of their own involvement in prostitution.

Part 1 provides an extremely detailed and thought-provoking look at the legal regulation of prostitution in Britain, then moves to examine academic discourse on prostitution. Joanna Phoenix structures this discussion around what she identifies as the "central and fundamental question:

in what ways are prostitute women like or unlike other women?" (p. 35). Rather than simply damn her predecessors for their failings, she concentrates upon how different explanatory models have opened up theoretical spaces and empirical questions. The same openness and generosity is evident in part 2 of the book, which explores how 21 British women understand their own experience of (predominantly street) prostitution.

Following in the best traditions of ethnography, Phoenix provides a richly textured account of these particular women's lives and relationships, an account in which interviewees appear as full human subjects, not simply as "prostitutes." A complex and sensitive portrait emerges, showing us women attempting to reconcile a sense of themselves as agents with an awareness of the constraints on action imposed both by structural and institutional factors (gendered labor markets, welfare and housing policies, prostitution law, etc.) and by individual psychobiography (personal histories of sexual violence, emotional abuse, neglect, etc.). The great strength of the book lies in its detailed description of the contradictions implicit in women's accounts of their own prostitution. Phoenix shows, for example, how interviewees simultaneously viewed their own prostitution as a strategic response to economic and material circumstances, and as "a trap that circumscribed their future possibilities and threatened their economic survival" (p. 115); how prostitution represented both a means of achieving independence from abusive male partners/parents/carers and an inevitable barrier to independence because becoming a prostitute "made them dependent on violent and abusive men for money and housing" (p. 125). Interviewees simultaneously identified as workers and as commodified bodies, as victims and as survivors. Prostitution was at the same time "nothing to do with sex and everything to do with sex" (p. 186).

Still more intriguing features of these women's accounts are, first, what Phoenix describes as the conflation of men and money (p. 135) and, second, the significance of race ideas for their lived experience and their attempts to make sense of it. All bar one of Phoenix's interviewees were white, eight had children of "mixed parentage," 19 had had black pimps, and racisms (both denigrating and exoticizing) were a "striking feature" of all their narratives—they stereotyped black men as dangerous and predatory, they also "talked about how 'attractive' and exciting black men were" (p. 173). This is not really incorporated into the book's overall analysis of prostitution, however. Why her interviewees "believed that their prostitute experiences were merely an accentuation of the paradoxical conflation of men, money and violence that shapes, structures and influences all heterosexual relationships" (p. 185), and why blackness was such a significant marker in terms of configuring the meanings attached to men, money and violence, are questions that the book leaves open.

This is indicative of the author's reluctance to move outside the subject world of her interviewees, and herein lies the only real weakness of the book. As E. P. Thompson remarked in relation to the historian's task (*The Poverty of Theory and Other Essays* [London: Merlin, 1978]), while

experience is a category that is indispensable, it is only "valid and effective . . . within determined limits: the farmer 'knows' his seasons, the sailor 'knows' his seas, but both may remain mystified about kingship and cosmology." Ultimately, Phoenix's focus on prostitute women's subjective experience seems to leave her as mystified as her interviewees about the systemic factors that underpin the contradictions they describe. Hence, at the end of the book, she finds herself unable to draw any conclusions about "specific policy interventions" (p. 189).

As a study of a small and not necessarily representative sample of prostitute women in Britain, the book is unlikely to help "towards displacing the 'either/or' analyses and debates that have bedevilled writing and research on prostitution" (p. 188). As thick description of those women's experiences, however, it is a study of outstanding depth and perspicacity that makes an important contribution to the body of empirical evidence on prostitution.

*Freaks Talk Back: Tabloid Talk Shows and Sexual Nonconformity.* By Joshua Gamson. Chicago: University of Chicago Press, 1998. Pp. x+288. $22.00.

Margaret Cerullo
*Hampshire College*

Three critical issues animate Joshua Gamson's exhaustively detailed study of queer presence on daytime TV talk shows. First, once we recognize that, and to some extent how and where, sexuality is socially constructed, how do sexual categories and meanings and the cultural authority embodied within them shift? Second, queer politics has been overtaken by an explosion of cultural visibility despite the regime of "don't ask, don't tell" that continues to define public policy. What institutional dynamics, then, structure that visibility and with what consequences for queer (lesbian, gay, bisexual, transgendered) citizenship, for full membership in the body politic? Third, where and when is the "public sphere" and how does it work in the mass-mediated, televisual culture that conditions our world and our worldviews and, in an irreducible mix, our possibilities for identity and identification and our rational assessments of political issues? Indeed, when and where is "the private," when private experience moves from property (derived from *propre,* one's own) to commodity, as happens on profit-driven commercial talk TV? The great strength of Gamson's book is to treat these issues together and to suggest that in the contemporary United States anxieties about and contest over the lines between public and private are intimately bound up with anxieties about and contest over the lines between normal and abnormal sexuality. Gamson argues that television, perhaps particularly talk television, is a crucial mediator of the ongoing cultural war over public participation and the ownership of public space.

In a strategy familiar from his last book, *Claims to Fame: Celebrity in Contemporary America* (University of California Press, 1994), Gamson undertakes his analysis through interviews with production staff, talk show guests, and focus groups of fans, as well as through participating in talk show studio audiences and watching and analyzing transcripts of every show in which there was gay, lesbian, bisexual, or transgender content in the years 1984–86 and 1994–95 (160 shows). This methodology enables him to describe what happens on talk TV and how it happens. He analyzes how talk shows shifted from the rational tolerant avuncular style of Donahue to the therapeutics of Oprah and Sally Jesse Raphael where "honesty is the best policy" and talk (their business) is the cure (1986–87) to the "staged democracy of troubled lives" that marked the 1993 Ricki Lake transformation of talk TV. Ricki Lake's goal and achievement were (in industry terminology) to "young" and diversify the audience for daytime talk television by going after 18–34 year olds from a broader demographic (though the more diverse 18–34 year olds remained 80% women, the long-standing audience for daytime TV). This was accomplished simultaneously by a near-exclusive focus on personal relationships and an accelerated pace with multiple guests, combined with new strategies of guest and audience recruitment via 800 numbers (rather than established issue-based organizations): "Know someone who hates her sister's boyfriend? Call 1-800-GO-RICKI." In the course of this transformation, gay/lesbian/bi/trans people went from an issue to a regular presence in the context of "relationship issues"—"I stole my best friend's boyfriend." At the same time, in this new production environment, conflict, performance, spectacle, and speed overwhelmed controversy, testimony, discussion, and content (p. 63).

Gamson argues that talk shows shift cultural authority in such a way that "bigots become freaks," often defining homophobia rather than homosexuality as the problem. Detached, "judgmental" homophobic "experts" are sabotaged by the authority of direct experience, and the ideology of "overcoming self-hatred" and being "true to one's self" is elevated within the talk shows' therapeutic ethos. He is quick to point out that the same move that normalizes homosexuality by naturalizing it operates within a dichotomous, biologizing logic of "either/or." Gender is also naturalized and rigidified at the expense of transsexuals, insofar as anatomy is still destiny and one's true self cannot depart from its biological ground. Bisexuals in this framework are rendered as inherently promiscuous (whatever their sexual practices). Their "failure to choose," to align with an either/or logic, violates the validation of monogamy that renders homosexuality tolerable. Thus, talk shows loosen the distinction between normal and abnormal sexuality, valorizing "honest" homosexuals, preferably "well-spoken," white, middle class (straight-appearing and discreet). At the same time, they redraw the lines of tolerance to exclude those whose presence they increasingly solicit and depend upon—"uninhibited," "unschooled," unorganized, unpredictable, multiracial, gender crossers and bisexuals. These are the rowdy and outrageous bull-dykes

and flaming drag queens and kings who keep middle (class) America, including middle-class American homosexuals, titillated—and increasingly uneasy. Talk shows thus, he argues, amplify the fault lines already present within queer politics: who can represent gayness, deeply cross cut by class, race, gender, and respectability. Moreover, the public visibility of sex/gender nonconformity, particularly the over-the-top variety preferred by TV talk and embodied in the lower and darker classes, heightens middle-class and straight anxiety that public space is out of (their) control, and moreover that inappropriate things are invading their living rooms, being "forced down their throats."

All the while Gamson is at pains to argue that it is the production requirements of commercial, for-profit TV entertainment rather than "any free floating hegemonic cultural imaginary" that structure all that talk television allows and disallows—in particular the ways in which it shakes up then reestablishes the stability of sexual and gender categories and boundaries in a seductive mix of threat and reassurance that is supposed to keep audiences riveted and willing to tune in again. It is almost as though talk shows (particularly those that feed on the "fringe" of transgender) mime, make visible, and enable the compelling and compulsory repetition of the unconscious of normalizing gender performance. They ensure the continual eruption of the sweet pangs of temptation and more open possibilities. "If gender can be made with surgery or clothing, if a woman can have a penis, can gender distinctions be natural occurrences of the body?" (p. 155). And talk shows encourage the reassuring and triumphal pleasure in their repression, via a punitive response to those who refuse "reality" ("You're a man, no matter what"). Perhaps, the "hegemonic cultural imaginary" informs commerce—the "production strategies" of talk TV—in a more intimate manner than Gamson allows.

*Interpreting the Self: Two Hundred Years of American Autobiography.* By Diane Bjorklund. Chicago: University of Chicago Press, 1998. Pp. xii+263. $25.00.

David R. Maines
*Oakland University*

Autobiographies are self-narratives in which authors try to present themselves to readers in a positive light. They are intrinsically moral documents that both draw from and reflect the cultural and social ethos of their time and thereby are forms of collective as well as personal narratives. It is in that fusion of the personal and the collective that Diane Bjorklund finds the material for sociological analysis. She sees two interrelated opportunities here. The first is historical analysis. How have autobiographies changed? What is the influence of context? The second is the analysis of the self. How are selves presented? What are the texts and vocabularies of the self? In these two sets of questions, Bjorklund ana-

lyzes "how people have made sense of their lives and experiences as they have answered the question, 'who am I'" (p. xi). Her analysis is not about selves and social processes, per se, but about their representations and how they are textually formatted. It is this very linguistic aspect of auto-biographies that allows for their sociological analysis, since autobiographers must use shared vocabularies of the self in order to present themselves to readers.

In the past 200 years, Bjorklund tells us, there have been over 11,000 autobiographies written in the United States. She has taken a sample of 110 of these documents, 80% by men and 20% by women authors, systematically distributed by decade from 1800 to 1980. She includes the famous and nonfamous and codes her materials for themes and story structure (analytical procedures are provided in two methodological appendices). She presents an array of appropriate theoretical considerations and then considers two main issues. The first, in chapter 2, pertains to the actual production of autobiographies—what she calls "autobiography as a social situation." The second (chaps. 3–6) presents the substance of the data analysis. Four models of the self are articulated here, and are located in the changing historical eras of American society. By organizing the book this way, Bjorklund can bring together her dual interests in vocabularies of the self and the changing historical contexts that shape those vocabularies.

Autobiographers are rather conscious rhetoricians; they are aware that "they are communicating with a future audience of readers" (pp. 16–17). Accordingly, they organize the events of their lives into a narrative structure that contains a plot or a central point that takes the potential audience into account. Bjorklund argues that autobiographers must configure three elements in their self-presentations to make them persuasive. The first is modesty: one must not be too vain or too humble. The second is honesty: readers must feel they are reading a truthful account. The third is interest: the account must secure the readers' attention. These elements are theorized as virtues that must be dealt with in terms of the shared understandings of them in the author's era. And, successful autobiographers deal with them deftly, since the three elements are always in potential conflict. Too much modesty may contribute to an uninteresting account; too much vanity may strike readers as untruthful; too much truthfulness (frankness) may violate norms of modesty; an account that is "too interesting" may stretch the boundaries of honesty. As the norms of credibility change, so do the configurations of these elements. They also change, Bjorklund writes, with "the changing assumptions about the self" (p. 42).

The first model of the self is what Bjorklund calls a morality play. The single largest category of autobiographers in the early 19th century was Protestant clergymen. They wrote religious autobiographies that focused on conversion, the struggle between good and evil, and the state of sin and wickedness. Conversions are attributed to God; the converted felt accepted and forgiven by God. Society, family, schools, and friends were

insignificant in these transformations of the self, and the idea of people having "true selves" was unheard of. Rather, the process of "knowing thyself" was to recognize the full extent of one's sinfulness and need for change.

The second model Bjorklund identifies is "masters of fate." In the second half of the 19th century, corresponding to the rise of industrialization in the United States, we see the self as something that can grow and develop, not merely change. The notion of "turning points" appears in the 1860s, one's childhood and education start to become acknowledged, and we find the businessman's autobiography as the self-made man in America. What was previously seen as sin now is seen as mere error, and overcoming error is not attributed to God but to "character." Through their character (hard work, discipline, thrift, trustworthiness), individuals can will themselves to become better people and to overcome their humble origins. At the personal level, these autobiographers told the story of progress that had been postulated by the social evolutionists.

The "uncertain self" is the third model and accompanies the development of psychology in the late 19th and early 20th centuries. Questions of motivation, the play of irrationality and the unconscious self, and emotional conflict become more prominent in autobiographies. "Finding oneself" now entails gaining insight into the essence of one's unique self; turning points are not developmental stages but epiphanies. Greater emphasis is placed on relationships—one's childhood and parents—and society is given credence as a source of external constraint. Autobiographers no longer needed to present themselves as exemplary people; rather, the focus was on depicting lives as lived. Autobiographies were also no longer univocally presented as fact but also as fiction. With the acknowledgment of imperfect memories and irrationality came the acknowledgment of imperfect autobiographical accounts.

Bjorklund's fourth model basically is a sociological one: the beleaguered self. Society is no longer a mere setting for the self but directly influences it. Issues of conformity arise, as do those of multiple selves. Greater emphasis is placed on education, intimate relationships, and group-linked selves found in current autobiographies focusing on ethnic identities. The vocabulary of "finding oneself" declines and that of "constructing oneself" appears.

This book is extremely interesting to read and is important for scholars working in the areas of historical sociology, social psychology, cultural analysis, sociology of knowledge, and narrative inquiry. Bjorklund's writing is clear and unadorned with verbiage, and readers will find her evenhanded analysis appealing. This book will stimulate curiosity and thought, but I suspect that most readers will find themselves a bit more cautious in their sense of certainty about social phenomena. Indeed, the "self" is not the same thing at different points in time; nor are things like society, education, family, emotions, relationships, religion, work, and all the other societal factors that sociologists study. Rather, if Bjorklund is right, they are changing configurations of salience and vocabularies of

living. And if that is correct, then Kenneth Gergen and others like him who argue social science can only articulate historically specific knowledge also are correct.

*A Genealogy of Manners: Transformations of Social Relations in France and England from the Fourteenth to the Eighteenth Century.* By Jorge Arditi. Chicago: University of Chicago Press, 1998. Pp. x+312. $35.00 (cloth); $17.00 (paper).

Peter Burke
*Emmanuel College, Cambridge*

Jorge Arditi's ambitious aim is to rewrite Norbert Elias with the help of Michel Foucault. He traverses the same ground as Elias did in his famous study of the civilizing process but offers an alternative model of phases in the history of manners and the relation between changes in manners and changes in what he calls "infrastructures of social relations" (it would have been useful if the author had explained the difference between this concept and Elias's idea of "figuration"). According to Arditi, manners in Europe, or Western Europe, or at least France and England—it is not clear to what extent France and England are supposed to represent a larger unit—went through three phases between the 14th and the 18th centuries, each "a function and instrument of the power of a different social group" (p. 221), first the church, next the monarchy, and finally the aristocracy. The church based its power on "revelation," the monarchy on "grace," and the aristocracy on "systemic relationality." Arditi therefore speaks of three "logics" and three "pragmatics."

   This model of change, which draws—but not uncritically—on the ideas of Foucault and of Pierre Bourdieu is an attractive one. It only remains to test it. To do this, Arditi concentrates his attention on a small group of texts, preferring, as he says, "depth" to "breadth" (p. 16). The texts include some of the most famous discussions of manners, from Erasmus's *De civilitate* (a key text for Elias), through Baldassare Castiglione's *Cortegiano,* Giovanni Della Casa's *Galateo,* and Nicolas Faret's *Honneste Homme,* to Lord Chesterfield's *Letters* to his son. However, equal attention is given on occasion to other texts, a somewhat miscellaneous group that includes plays and personal documents such as William Shakespeare's *Henry IV,* Pierre Corneille's *Cid,* Samuel Pepys's *Diary,* and the duke of Saint-Simon's *Mémoires.* The assumption behind the choices (and the lack of any serious attempt to justify the choices) appears to be that of the classic form of *Geistesgeschichte* practiced in Central Europe in the first half of the 20th century by scholars such as Erich Auerbach, that is, that all texts produced in a given period reflect the "spirit" of that period, the *Zeitgeist.* There is no room in this model for cultural conflict, and Arditi has remarkably little to say about the conflicts over manners that took place in early modern England and France,

notably between the rival styles of "court" and "country," Catholic and Protestant. On occasion, Arditi draws on the history of language for evidence, as in the case of his interesting remarks about the rise of the term "etiquette." It would have been still better if he had done this more systematically, examining the changing place of such keywords as *chevalerie, courtoisie, honnêteté, galanterie, politesse,* and so on within the linguistic field of good and bad manners.

There are other serious flaws in this study. In the first place, Arditi does not seem to have read some of the most relevant secondary literature in his chosen area. A second major flaw in this study is that it ignores microsociology, especially the sites in which good manners were taught or put on display. The vague term "church," sometimes equated with religion, is never broken down into constituents such as bishops, parish priests, monks, or nuns. There is no discussion of the court as a locale, whether it was mobile or fixed, or dominated by warriors or civilians, clergy or laity, or men or women (the relation between manners and gender is a topic virtually ignored in this book).

A third flaw is perhaps the most serious of all. Since church (or churches) monarchy and aristocracy existed in France and England throughout the period (apart from the hiatus of the Commonwealth and Protectorate, 1649–60), the relative importance of the three elements (or as Arditi calls them, "groups") ought to be discussed in each section. To confine the church to the first section, the monarchy to the second, and the aristocracy to the third risks producing a purely circular argument. In the case of the 17th century, for example, historians disagree about the importance of the monarchy relative to the other two groups. Their disagreements are of great relevance for Arditi's thesis, yet he chooses to ignore them. He also ignores—like Elias—the possible importance of a fourth player in the game, the city, although the importance of urban culture in the history of manners is well known and has often been discussed by historians. This book cannot be considered a success either as sociology or as cultural history.

*Riding the Black Ship: Japan and Tokyo Disneyland.* By Aviad E. Raz. Cambridge, Mass.: Harvard University Asia Center (Harvard East Asian Monographs 173), 1999. Pp. xiv+240. $40.00 (cloth); $19.95 (paper).

Gary Alan Fine
*Northwestern University*

Americans often seem to enjoy nothing so much as pouting and fretting about the sorry state of their popular culture. Whether it be McDonalds, Tupac Shakur, Las Vegas, *Pulp Fiction,* Marilyn Manson, or the Disney empire, our pleasures mark us as the boorish parvenus that we secretly know ourselves to be. Such a view depends mightily on being blissfully unaware of the leisure choices of other cultural groups and how others

appropriate our own culture. As a result, Americans often compare themselves unfavorably to the Japanese—holding tightly to a stereotype of Japanese popular culture as less violent, less commercial, and less sexual. Whatever virtues one can point to in Japanese society, it is hard to attribute them to the bracing morality of their popular culture choices.

Aviad Raz, an Israeli sociologist, has done American students of popular culture a signal service by examining the popularity of Tokyo Disneyland (TDL) in Japanese culture. TDL has an attendance of approximately 16 million per year, compared to 12 million at Disney World's Magic Kingdom. For those relatively unfamiliar with the nuances of Japanese society, this fine work of comparative ethnographic scholarship sparkles with insight in how cultural translation operates.

Much of the writing of Tokyo Disneyland has assumed that it is a "black ship" representing the imposition of American culture on Japan. This powerful image, operating on several levels, refers to the black ships of Commodore Matthew Perry that sailed into Tokyo Bay in 1853, "opening" Japan to the west. Raz's contention is that such a view of cultural hegemony is fundamentally misleading. Rarely does one society simply impose a cultural domain on another without that domain being shaped in the process. Cultural trade is negotiated in practice. Raz demonstrates that the Japanese take American culture and transform it to meet their needs, blending and recombining. While the structure may seem Americanized, the way that structure is reconfigured is distinctively Japanese. The manuals that the American parent company provides to their Japanese partners are only occasionally relied upon; the Japanese company running the park is at least as influential in explaining how workers (cast members) should treat customers. Both the entrance to TDL, which has been transformed from "Main Street" to "Meet the World" (altering Japanese history to fit the demands of corporate Japan), and a distinctive Japanese Cinderella's castle represent the needs of the audience. In the castle, Japanese visitors are instructed in the "meanings" of European fairy tales, transforming those meanings in such a way that the Japanese can appreciate their "child-like magic." These examples, and others, demonstrate that what appears to have been taken directly from the Disney corporate model has been "oriented" for the Japanese.

For employees, this shaping takes the form of the stories that coworkers share. These narratives, recognizably similar to contemporary legends or rumors, provide explanations and moral claims about the conditions of work. One of the most popular of these (in both Japan and the United States) is the "popcorn story." In this text, a costumed cast member sees a young child spill a box of popcorn. The employee scoops it up, refills it, and hands it back to the screaming child, without breaking stride. This tale, involving the creation of a "magic moment," simultaneously enshrines the individual worker and justifies the corporate ideology. While not all stories are positive, each attempts to tame the experience of working by building community.

*Riding the Black Ship* is particularly impressive in that Raz is not satis-

fied with a single perspective to understand the Japanese transformation of this amusement park. Integrating previous examples of "Disney studies," Raz examines three components of the meaningful character of TDL: (1) the frontstage work of communicating the text to an audience; (2) the backstage in which workers are shaped and resist the organization's ideology; and (3) the cultural critique implicit in TDL. The range of analyses makes this book stand in contrast to more narrow social scientific examinations of the Disney phenomena.

For better and worse, Raz, neither a Disney-phile or Disney-phobe, is measured in his analysis. Rather than emphasizing the power of Japanese resistance to the Disney ideals or the power of Disney to shape Japanese culture, he emphasizes how the Japanese, in both small details and larger choices, take Disney and makes it "Japanese." This strategy is an intellectually defensible one, although it has the effect, as does much "moderate" and "prudent" analysis, of leaving the reader unsure the extent to which the phenomena discussed occur more broadly. We learn that when cultures come into conflict that a merging will occur, but how and how much remain open questions, presumably a result of local conditions.

One area that could have been expanded is how emotions are massaged in corporate entertainment, both the emotions of the guests and the emotions of the employees. It is evident that Americans and Japanese have different patterns of emotional management and display. Although Raz touches on this issue, the discussion of the differences would have been more compelling with an extended treatment of emotional variation, drawing on the now extensive literature on the sociology of emotions. Do guests respond differently to the kinesthetic experiences of particular rides (roller coasters, haunted houses)? Does cynicism emerge in the same organizational spaces? On a more general level, could the way that cultural elements are filtered through experience explain how seemingly similar forms of popular culture affect societies differently?

*Riding the Black Ship* achieves the goal of powerful ethnography. It permits us to visit a shadowy world of which we are only partly familiar and provides both the insights of the participants and the insights of the marginal observer. Readers of this fascinating volume can compare their own experiences with an American Disney World to how the Japanese— workers and guests—experience the park. This study reminds us just how similar and how very different "the same" cultural form can be when enacted in two distinct social worlds.

*Representing: Hip Hop Culture and the Production of Black Cinema.* By S. Craig Watkins. Chicago: University of Chicago Press, 1998. Pp. xlv+317. $30.00.

Jon Cruz
*University of California, Santa Barbara*

Craig Watkins presents a compelling assessment of how recent transformations in American culture have altered the mode as well as the manner in which black youth culture has been represented over the last two decades. Since the 1980s, a number of important developments have occurred, each of which has been brought into a somewhat inadvertent relationship, enabling what Watkins refers to as an explosion of cultural innovation and production among black youths. The conservative populism of the 1980s championed a host of crisis discourses, many of which centered on race. These discourses reinvigorated the notion of black youths as constituting one of the most symbolically charged social problems, treating them as symptomatic of the worst deviations from traditional-inflected arguments over "family" and "personal responsibility." Those most marginalized came back into the symbolic center of social scientific concern.

While social scientists rekindled their talk of the underclass, the television industry, true to the fiscal lure of the lurid, focused on the presumably more sales-worthy but seamy side of things. They gave us frequently racialized "reality programming." Consider the important splash made by the TV networks a decade ago when reality programming arose as a major and relatively inexpensive television product. Recall, for example, how CBS launched the program *48 Hours* with veteran network anchorman Dan Rather leaving his desk and the safety of the studio to ride, with cameras running, in a squad car that toured crank-infested ghetto neighborhoods. Rather had joined the already growing company. Popular television culture was already featuring programs like *COPS* and *Hard Copy.* The genre of street and community policing continues with programs like *Real Stories of the Highway Patrol.* Such programs have been consistent in their abundant portrayals of policing America's most unwanted at those knotty intersections where the urban and the racial come entwined, where those presumably naughty by nature are most frequently found. What seized front stage was televised moral panic, frames of deviance with marauding individuals out of control, and frequently beyond the reach of the law and order. With the exception of a few laugh-tracked sit-coms, little else pertaining to African-American, Hispanic, and other minority communities came into view.

At the same time, black youths responded to both the negative populist indictments as well as the structural forces that hemmed them in by forging new voices and launching new cultural practices. Of these, hip hop culture has proved to have some staying power. Simultaneous changes in the political economy of cultural production added to the mix. Music

Television, or MTV, notorious for its attempt to screen out all black music, soon had to succumb to the tremendous popularity of rap and hip hop music. By the early 1990s, hip hop culture had become one of the economic staples of the music industry. Black cultural developments involving dance, poetry, and music making converged in new expressive forms. Watkins argues that the new developments within black youth culture demonstrate how the latter are not at all fathomed by the neoconservative caricatures of social dysfunction. Rather, black youth culture has become the most important social context for young people to speak to the social fate that has structured their lives. Watkins shows how black youth culture displays young black persons being present in their own making. It is, in part, a study that aims to restore agency to the most economically and politically marginalized.

A good deal of the book pivots its analysis upon the rise and success of black filmmaker Spike Lee. Watkins analyzes Lee's films as cultural products that mediate between neoconservative populism and the significant development of "ghettocentric" black filmmaking. In this regard, Watkins sees Lee's work as symbolizing an important cultural, political, and economic nexus. Watkins charts the recent history of black filmmaking against which he situates Lee's emergence. But rather than treat Lee's films as autonomous developments, Watkins shows with convincing care how they actually inhabit the cultural space created by an already socially, culturally, and economically insurgent as well as politicized hip hop culture. Lee's films confront much of the contested imagery of race relations at century's end. Watkins shows just how skillful Lee is in capturing and engaging the tension-ridden cultural landscape that hosts both the racial discourses that have demonized black youth as well as the new public explosion of black youth struggling over representations of them and, in turn, by them.

*Representing* is a nicely crafted and well-written book. While many studies have addressed the cultural politics of the last few decades and have examined a wide range of popular culture, Watkin's study is valuable for his deft integration of multiple arenas. Most valuable is his ability to link black youth culture to the recent success of black cinematography, a move that shows how cultural analysts might conceptualize the difficulties of bridging cultural practices, political economy, institutional transformations, and social problems discourse.

*Whiteness of a Different Color: European Immigrants and the Alchemy of Race.* By Matthew Frye Jacobson. Cambridge, Mass.: Harvard University Press, 1998. Pp. 338. $29.95.

Stephen Steinberg
*Queens College and Graduate School, CUNY*

A visitor from Mars to the annual meetings of the American Sociological Association would have observed that circa 1990 race scholars assumed an odd forensic pose. As they spoke from the podium, they would periodically raise both hands and wiggle two fingers. The reason is found in the opening sentence of *Whiteness of a Different Color:* "It has become customary among academics to set words like 'race,' 'races,' 'Anglo-Saxons,' or 'whiteness' in undermining quotation marks" (p. ix).

The logic for bracketing all referents to race in quotation marks is perhaps self-evident. If races are not *real,* if they are not natural categories or rooted in biology, then why should social scientists employ the language of racism, and speak about race as though it were real? Are we not giving credence to the very assumptions that we ought to be debunking?

The thrust of *Whiteness of a Different Color* is to trace the "racial odyssey" of immigrants from Eastern and Southern Europe who were at first regarded as "racial other," and relegated to the twilight zone between "black" and "white." We are speaking here of groups who in early 20th-century vernacular were classified as "Hebrews," "Celts," "Mediterraneans," "Iberics," "Slavs," "Teutons," and the like. Not only did these groups have obvious reason to disassociate themselves from blacks, but they played an active role in forging a monolithic whiteness that included them, and in doing so, furthered the objectification and isolation of blacks. The historical and moral lesson is that these "in-between peoples" were directly implicated in the reconstruction and validation of America's racist lexicon. The analytical lesson is that "race" is not immutable or fixed in biology, but rather something that is fluid, historically relative, and socially constructed.

Chronicling how these racial pariahs came to be "white" lies at the center of Jacobson's intellectual project. Part 1 provides a political history of whiteness, beginning with an act of Congress in 1790 that limited naturalization to "free white persons." This unitary conception of whiteness was shattered by the arrival of the "Celts" during the mid-19th century, and of "Hebrews," "Italians," "Slavs," and others at the end of the century. Now the racial divide was drawn between "Anglo-Saxons" and these other European "races"—36 of them according to the 1910 Dillingham Commission. Thus, for Jacobson, "the history of American nativism from the 1840s to the 1920s is largely the history of a fundamental revision of whiteness itself" (p. 68).

A chapter on "Becoming Caucasian, 1924–1965" then traces the process through which these groups became incorporated under the ex-

panded white umbrella. Jacobson shows how the escalating civil rights struggle sharpened the binary distinction between black and white. The end result was that race lost its salience when it came to southern and eastern Europeans, culminating in 1965 with the abolition of the national origins quotas that had governed immigration policy since 1924. Increasingly, race referred to "the longstanding, simpler black-white dyad of the Jim Crow South" (p. 117).

Other chapters probe the fluidity of race by focusing on particular historical moments (1877) and groups (Jews), on the role that empire played in forging the logic of "pan-white supremacy," and on the politicization of "white ethnicity" in response to the escalating Civil Rights movement. The chapters in this book do not form a seamless whole, but all of them provide valuable historical documentation on the evolution of "whiteness" as we know it today.

The hallmark of the new genre of whiteness studies is to turn the racial lens on whites, illuminating the subtle and nefarious ways in which whites fabricated a racial identity that served their self-interests at the expense of nonwhites. This is a welcome theoretical development. Yet in his zeal to prove that races are not real, Jacobson gives short shrift to that great sociological adage that if people believe something is real, it is real in its consequences. As Lorraine Hansberry wrote in *Les Blancs,* racism may be only a "device" for domination, but a man "who is shot in Zatembe or Mississippi because he is black is suffering the utter *reality* of the device. And it is pointless to pretend that it doesn't *exist*—merely because it is a *lie!*" (*Les Blancs: The Collected Last Plays of Lorraine Hansberry,* edited by Robert Nemiroff [Vintage, 1973], p. 122).

*Crosscurrents: West Indian Immigrants and Race.* By Milton Vickerman. New York: Oxford University Press, 1999. Pp. xi+211. $19.95 (paper).

Vilna Bashi
*Northwestern University*

Milton Vickerman, like every ambitious author, proffers a text that succeeds at many but not all of his endeavors. If you seek an ethnographic work rich in interview material that fully describes the on-the-ground experience of race and ethnicity for West Indians in New York, *Crosscurrents* is it. However, *Crosscurrents* is less successful in providing a strong theoretical framework by which to understand race, ethnicity, and identity in the United States.

Black immigrants face a unique set of social circumstances upon entering our borders. Vickerman borrows from Bryce-Laporte (1972, "Black Immigrants, the Experience of Invisibility," *Journal of Black Studies* 3 [1]: 29–56) the term "cross-pressures" to name the contradictory circumstances that mark the West Indian experience in the United States. "In

the case of West Indians, the primary forces producing cross-pressures are, on the one hand, conservative socialization in societies in which they form majorities and which urge them to downplay race while emphasizing merit; and, on the other, immersion into a host-society (the United States) which provides greatly expanded opportunities for upward mobility but which also views race as being a key organizing principle" (p. 5).

The book has many strengths. Vickerman compiles an impressive array of important and relevant literature on the economics of migration and provides a strong explanation of how race works in Jamaica (which he later contrasts with the United States). The author offers us a typology of coping strategies that categorize West Indian responses to different kinds of racial encounters. Succinctly summarized in a 2×2 table, these action-reaction pairs map the statics, but Vickerman covers the dynamics of responses to racism as he explains that West Indians develop a new racial consciousness over time. He also chronicles the varying attitudes they may have toward African-Americans, a controversial topic for which we have been eager for more writings.

The contribution Vickerman makes with this much-needed chronicle of the West Indian racial experience in New York is, in the main, descriptive. It would have been helpful to see more formal theory applied to the empirical data in order to provide us a more generalized knowledge about race, identity, and ethnicity. Vickerman's analysis rests on a shifting theoretical frame that can only be discerned from the text. At times, Vickerman seems to suggest that ethnic and racial categories are somewhat mutually exclusive, as if one cannot have both an ethnicity and a race at the same time. (Other writers on the immigrant and second generation have suggested this too, although I argue this framework is confusing at best—see Bashi, "Racial Categories Matter because Racial Hierarchies Matter: A Commentary," *Ethnic and Racial Studies* 21 [September 1998]: 959–68). For example, Vickerman suggests at the book's end that three options exist for West Indians: they may choose to identify as "black," or instead as "West Indians," or choose a pattern of "identifying with African Americans on some issues but distancing from them on others," as if West Indians do not do all three simultaneously. Although he writes that the "norm" for West Indians will likely be a "complex pattern of identity formation" (p. 168; i.e., that simultaneity is possible), he also writes that since West Indians "share the fate of many African Americans, it could be said that they are forced by the society to become 'black'" (p. 169; i.e., instead of their other racial and ethnic options). He sometimes uses the terms ethnicity and race interchangeably, and at other times, he treats ethnicity as if it is the broader category that encompasses racial ones. Confusing language sometimes results from the absence of discrete handling of racial and ethnic categories. For example, he must create a new ethnoracial category, "panblack," to describe moments when black subgroups come together, since he treats "black" as an ethnic

choice, that is, as if it is not inclusive of all who share black phenotype. He also uses the phrase "entrenched ethnic hierarchy" when "racial hierarchy" seems to be more appropriate to the U.S. context. Although not clearly specified by the author, the theoretical framework on which the work rests shapes Vickerman's conclusions—conclusions that might be differently stated had he used a framework that sees racial and ethnic categories as distinct, where race is hierarchical and ethnicity is not (Omi and Winant, *Racial Formation in the United States: From the 1960s to the 1990s*, 2d ed. [Routledge, 1994]; Cornell and Hartmann, *Ethnicity and Race: Making Identities in a Changing World* [Pine Forge Press, 1998]), or one that treats ethnic categories as a subset of racial categories (Smedley, *Race in North America: Origin and Evolution of a Worldview* [Westview Press, 1993]; Bashi and McDaniel, "A Theory of Immigration and Racial Stratification," *Journal of Black Studies* 27 [May 1997] 5: 668–82), and not the other way around.

With data from 106 interviews with Jamaican men, Vickerman's book is a significant compilation of new fieldwork among West Indians in New York. Its ethnographic richness, and the important contribution he makes with his typological organization of the data, ensures that this work will be cited in studies on West Indians, race, and ethnicity, for years to come.

*Help or Hindrance? The Economic Implications of Immigration for African Americans.* Edited by Daniel S. Hamermesh and Frank D. Bean. New York: Russell Sage Foundation, 1998. Pp. x+314. $45.00.

Lisa Catanzarite
*University of California, San Diego*

A thorny and unresolved question in the ongoing debate over immigration is whether and how the latest wave of newcomers may harm African-Americans. Hamermesh and Bean's volume represents the first concerted effort to assess this issue, bringing together state-of-the-art research by economists on fiscal consequences of immigration for native blacks and enlisting outside assessments of these studies in several concluding chapters.

Labor market analyses comprise the book's backbone. Two theoretical chapters consider substitution of capital for various types of labor: Johnson's simulation takes into account consequences of job segregation and concludes that low-skilled immigration entails a small income transfer from (disproportionately minority) lower-skilled workers to their higher-skilled counterparts. Borjas suggests greater minority-white differences, arguing that benefits of immigration accrue disproportionately to whites—whose ownership of capital far exceeds that of blacks—while costs fall most heavily on less-skilled workers; African-Americans stand to lose from *both* high- and low-skilled immigration, since owners of capi-

tal benefit from the former, and low-skilled labor competes with the latter.

Did immigration inhibit black capital ownership? Fairlie and Meyer discern no adverse impact of immigration (or disaggregated Asian immigration) on African-American self-employment across metropolitan areas.

Hamermesh examines the thesis that immigrants take jobs natives do not want and finds immigrants no more likely than native, non-Hispanic whites and *less* likely than blacks to work late hours or in hazardous industries: he concludes that immigrants do not compete directly with blacks, but they do compete with whites. While he tests whether European-origin immigrants diverge from other foreign-born immigrants, one wonders if results would differ were non-Europeans distinguished by some approximation of skill (e.g., educational attainment, ethnic origin).

Reimers's and Butchers's metropolitan area–level analyses of 1980–90 immigration changes do consider skill. Reimers demonstrates the sizable negative effects of increased immigration on men's hourly wages among better-paid, native white and black high school dropouts, but an anomalous, *positive* effect for Mexican-American dropouts. By contrast, Butcher detects no influence of increased immigration (or less-skilled immigration) on changes in black (women's or men's) hourly earnings, annual earnings, employment, or weeks worked but detects a positive impact on the black-white *gap* in men's annual earnings. While the analyses are quite rigorous, there is little attempt to explain why some outcomes should be affected by immigration and why others are not.

How might immigration bear on prelabor market conditions? Betts provides convincing evidence that immigration inhibits African-American (and, in California, native Latino) high school completion, arguing that competition for school resources increases the marginal cost of education to native minorities. Similarly, Hoxby's study suggests that, in moderately selective colleges, competition for resources crowds out native blacks and Latinos. In more selective institutions, crowding is ethnicity specific: schools apparently meet affirmative action targets by preferring immigrants of higher socioeconomic status over disadvantaged, native coethnics.

Evidence on other prelabor market outcomes is mixed. Grogger's analyses of the National Longitudinal Survey of Youth and census data unearth no influence of immigration on crime among young black men. Zax examines the import of immigration (without regard to ethnicity or skill) for neighborhood effects among blacks, finding ambiguous impacts of short-term immigration but that long-term immigration increases blacks' segregation and reduces their incomes.

Hence, immigration yields some fairly sizable, negative consequences for African-Americans, particularly in educational institutions and the labor market. However, some labor market outcomes appear unaffected. Definitive answers in this field apparently remain elusive, yet this book begs for a concluding chapter by the editors, speculating on consistent and divergent findings and discussing implications. (The editors might

also have requested more uniformity across chapters, e.g., attention to low-skilled immigration, relegation of less-important estimations to footnotes, and elimination of $p$-values $< .10$.)

The editors instead invite outside experts' assessments in several concluding chapters, providing useful discussion and raising critical insights, some which overlap with my own primary concerns. First, if the aim of the volume is to provide a comprehensive set of analyses of economic implications of immigration for African-Americans, then at least one substantive chapter devoted to welfare would have been warranted. (As Shuck argues, support for the 1996 welfare reform surely was shaped by public perceptions regarding immigration.) Second, most of the empirical chapters provide rigorous analyses comparing metropolitan areas. While their methodological and substantive contributions are to be applauded, I agree with Tienda that the literature would benefit from a move away from "area" studies. Greater attention to analyses of (a) particular local labor markets with heavy immigrant inflows and (b) occupational segregation (including job-specific demand for immigrant vs. native minority workers and wage consequences of segregation) would yield new insights on mechanisms by which the new immigration might differentially affect African-Americans, other native minorities, native whites, and earlier immigrants.

*The Bottom Rung: African American Family Life on Southern Farms.* By Stewart E. Tolnay. Urbana: University of Illinois Press, 1999. Pp. 243. $45.00 (cloth); $19.95 (paper).

Steven Ruggles
*University of Minnesota*

This exciting and innovative analysis should be required reading for anyone seeking to understand the origins of African-American family structure. Tolnay focuses on the Southern farm population in the first half of the 20th century, a group that in 1910 included 46% of African-Americans. The bulk of the study is based on samples of the decennial censuses of 1910 and 1940, supplemented with qualitative material drawn from the Federal Writer's Project and other sources. The first six chapters examine work and schooling, marriage formation, fertility, marital dissolution, and geographic mobility of African-American farm families in the South. On each of these topics, Tolnay provides an overview of trends and race differentials together with an admirably clear and concise discussion of alternate theoretical interpretations. The readability of the book extends to the quantitative analysis: with the exception of a few logistic regressions, it consists of simple and elegant descriptive statistics. Although the book has much to offer specialists, it is suitable for a broad audience, including undergraduates in introductory courses.

The history of the African-American family has been the most contro-

versial branch of family history over the past three decades. Tolnay's carefully reasoned and superbly documented analysis is not entirely dispassionate, but it is a more balanced interpretation than is usual in the subfield. Nevertheless, in a few instances, Tolnay's interpretation of the Southern farm family will be controversial. For example, Tolnay takes issue with the thesis proposed by Antonio McDaniel and others that differences between African-American and Euro-American family composition can be ascribed to African cultural inheritance or the experience of slavery. Tolnay points out that racial differences in family structure actually grew over time, instead of diminishing as one would expect if they were simply a carryover of the past.

In the last two chapters, the focus shifts from the first to the second half of the 20th century and from Southern farms to Northern and Western cities. Tolnay demonstrates convincingly that the sharp increase in African-American unmarried fertility and marital instability after the war cannot be ascribed to northward migration. The changes in family structure occurred in both North and South, and in all periods migrants from the South were more likely to reside in two-parent families than were persons born in the North.

What, then, caused the postwar shift in African-American family structure? For the South, Tolnay points to the displacement of sharecropping by agricultural mechanization, which undermined the family economy. In the Northern cities, Tolnay stresses the impact of residential segregation together with limited economic opportunity for young African-American men in central cities. At the same time, however, Tolnay suggests that the Aid to Families with Dependent Children (AFDC) program also contributed to unmarried fertility among blacks. Tolnay acknowledges that there is a sizeable body of research showing that state differentials in AFDC benefit levels show only a weak or insignificant association between AFDC and unmarried motherhood, and that there is a poor fit between trends in AFDC benefit levels or participation and African-American unmarried fertility. He argues that the results might be different, however, if one focused on the population most at risk of unmarried childbirth. Be that as it may, the evidence in those studies is persuasive that, in the aggregate, changes in the AFDC program cannot explain a substantial proportion of change in family structure, and Tolnay offers no data to the contrary.

Despite its many merits, Tolnay's book has a few limitations. The analysis of Southern African-American families in the first half of this century—which is the heart of the book—seems to me too narrowly focused. Tolnay examines exclusively those African-Americans residing on farms, but even in 1910, 51.6% of Southern African-Americans did not reside on farms. Nonfarm employment was an important part of the African-American experience even under slavery. Moreover, the boundaries between farm and nonfarm residence were fluid; individuals and families might reside on the farm during some parts of the life course and not in others. A second limitation is chronological. The data exist to carry the

analysis of Southern families back to Reconstruction. Since so much of Tolnay's interpretation rests on chronological change, I was disappointed that the prewar analysis is based entirely on two data points. I have a related complaint about Tolnay's analysis of the effects of northward migration on the family. One of Tolnay's purposes was to show that writers such as E. Franklin Frazier and Charles J. Johnson were wrong about the impact of migration from the South on northern African-American family structure. Frazier and Johnson, however, were writing in the 1930s, whereas Tolnay's migration analysis begins in 1940. To test those interpretations effectively, he would have had to assess the impact of the earlier waves of migration.

Despite these quibbles, this book is a major contribution to the literature. With the exception of the point about AFDC, Tolnay's arguments are uniformly persuasive. His analyses of the effects of African cultural heritage and of the effects of northward migration after 1940 are especially cogent. Tolnay's lucid discussions of alternate theoretical interpretations together with his accessible presentation of new historical evidence will clarify the debates on the African-American family.

*Cottage Country in Transition: A Social Geography of Change and Contention in the Rural-Recreational Countryside.* By Greg Halseth. Montreal: McGill-Queen's University Press, 1998. Pp. xvii+283. $60.00.

James R. Hudson
*Center for Excellence in Association Leadership*

Social change never occurs without some stress. Halseth adds to the evidence of this axiom with a careful analysis of a rather unique situation. Around a number of Canadian lakeshores, there has been built over the years a large concentration of cottages, which were initially used for summer vacations. There were idyllic myths and images associated with the lakeside living of quiet, family-oriented summer vacations. The image obscured several latent problems, such as pollution, few services, lake access, road maintenance, and a clear class distinction between the rural residents and the cottagers. These problems became exacerbated as many cottagers began to convert their summer homes into residences to be used all year. The consequences of this process provide the focus of *Cottage Country in Transition.*

Halseth selected two lakeside cottage communities for comparative analysis and employed a wide range of methods. The data from these methods are very comprehensive and reported with a good deal of care and caution, and the amount of detail in these data permitted an exhaustive examination of a large range of issues. There are 42 tables and 32 figures, which gives some measure of the data collected and analyzed. Because of the detail in these data, Halseth is able to fine tune his analy-

sis, capturing subtle nuances and differences between and with these communities and the rural residential areas.

While it is not easy to summarize all of these data, there are a number of findings that support Halseth's central thesis of change and contention in the cottage country. The framework Halseth employs to analyze these data is modified political economy, and he emphasizes that in a complex society like the one presented here, consideration should be given to the economic, social, and political arenas. Each has its own dynamics, but taken together, they provide a more complete explanation of the data than if considered singularly.

The two distinct populations living in these rural recreation areas are the rural farm population, scattered across a wide geographic area, and the cottagers, clustered along the lakeshores. The rural farm population has long-established roots in these areas and have developed over time a number of social institutions that have served them well. They enjoy modest incomes and education attainment. Many cottagers, too, have long-established roots, and some of the cottages have been in the same family for more than one generation. The cottagers have substantial incomes and university-level educations.

The points of contention between these two populations began to become more marked as the conversion from summer cottages to year-round residences gained momentum. One of the reasons for the increase in conversions was that the cottagers began to use their summer houses for their retirement. As they became year-round residents, they wanted increased services, better road maintenance, which included having private roads made public and protection from pollution of the lakes, which would insure the value of their property. To gain these ends, the cottagers created rate-payer associations. The rural farm residents had long used the townships as their means for collective action. In one study area, Halseth documents that the cottage associations were much better organized than the townships, more articulate, and had deeper pools of talent in their community to draw upon in mobilizing for political and social action. Indeed, these rate-payer associations were given certain legitimacy at other government levels when information and consultation was deemed appropriate for planning purposes.

While the major focus of this study is on the local level, the problems of rural versus recreational has not escaped the attention of other governmental levels. Issues such as pollution, regulated growth, and increasing tourism have given them an interest in what is happening to the cottaging areas. For the local residents, both cottagers and the rural farm communities, the possible incursion of other levels of government raised concerns over local control.

One of the most intriguing aspects of this book is the amount of research that has been done in Canada in the area of social geography and how little of this has found its way into the research in the United States. Halseth, for his part, cites very little work done in the United States in

building his theoretical and analytic framework. While this does not detract from a very sophisticated and interesting study, the richness of his work and his contribution to our understanding of social change might have been enhanced had he noted some of the parallels between the conflicts analyzed in some of the gentrification studies in the United States and those in the cottage communities in Canada.

*Networks in the Global Village: Life in Contemporary Communities.* Edited by Barry Wellman. Boulder, Colo.: Westview Press, 1999. Pp. xxiv+377.

Larissa Adler Lomnitz
*National University of Mexico*

Network analysis has a substantive history done both qualitatively and quantitatively by anthropologists, historians, and sociologists. This book represents an important addition to those studies, as it offers an innovative idea about the nature of contemporary communities, complemented by contributions to methodologies that can be used in the studies of networks, and a very well-chosen set of case studies in different sociocultural systems. Community is conceived as a set of personal, *ego-centered* social networks rather than a *space* where social ties are reproduced, as it was traditionally defined. This change in perspective allows researchers to find people's supportive community ties with kin, friends, and relatives no matter where they live, particularly in societies where technology allows people to keep in touch virtually, if not actually, on a daily basis. According to Barry Wellman, the editor and the author of several of the papers, this theoretical contribution of what community is today has been the "greatest victory" of sociology in the last decades.

The book contains 11 chapters ranging from developed and underdeveloped capitalist societies to communist and postcommunist ones. This excellent choice of case studies shows that although social exchange networks are universal, the cultural characteristics of each society determines the criteria by which individuals establish, maintain, and use ties.

In the first two chapters by Wellman and Potter, the authors provide a battery of concepts and techniques that allow researchers to move from a loose description to a systematic analysis of networks. This is particularly clear in the analysis of personal communities where a network is defined by the ties of interests that each has with the focal person. Different variables define the type of networks, and as a result, individuals maintain a menu of support networks. The empirical reference of these papers was a study in greater Toronto.

Another interesting methodological and theoretical contribution is Lee and Campbell's study of neighbors networks of black and white Americans in which the authors show that the urban neighborhood community of African-Americans appear to be a more intensive spatially concen-

trated phenomenon than it is for whites. The same is true with Ferrand, Mounier, and Degenner's study of the diversity of interpersonal ties and social positions in France. The findings of these papers are based on survey data that creatively combine variables, which allow us to understand the nature of networks and to generalize to the larger society.

I found Espinoza's paper to be of particular interest, as it poses the question of the survival of the urban poor within the description of economic and political changes that have taken place in Chile since the introduction of neoliberal reforms. Within that frame and the empirical study of family survival strategies presented, the author concludes that the economy of survival depends more on managing social ties to gain access to resources than on the strategic organization of goods and services. This relates to Wellman's idea of individuals having menus of different social networks that can be activated as needed.

Sik and Wellman's paper on network capital in capitalist, communist, and postcommunist societies argues that although people and organizations use networks everywhere, they use them more in communist and postcommunist societies. The authors describe the use of networks in the three systems, introducing the idea that network capital superimposes state and market more in postcommunist societies as uncertainty is added to the scarcity and inefficiency that are characteristic of communist societies.

The next three papers deal with Asian situations: the "institution of *Guanxi*" in China that shows how the Chinese in the communist regime use their web of *Guanxi* to get jobs. Otani's paper studies personal community networks in contemporary Japan concluding that although traditional forms of personal networks still provide the base structure for Japanese culture, there is a growing trend of individualism in Japanese society. Salaff, Wong, and Siu-lun contribute a paper about using networks to exit Hong-Kong in which they combine quantitative and qualitative data. Their sample includes working, lower-middle and upper-middle classes intending to leave Hong-Kong and to migrate to North America. The type of support needed in each case differs according to the resources of each group.

In conclusion, this is a useful collection of papers for researchers interested in community studies and social networks. It clearly introduces the concept of multiple personal networks as the basis of communities. The fact that formal social network analysis was developed in North America, where settlement patterns are often disseminated and communicated by new technologies, has allowed North American social scientists to discover this pattern of the personal network community. However, the case studies presented in the book, embedded in different types of societies whose history, urban patterns of development, and cultures differ from the North American ones, show that locally based communities have not necessarily disappeared everywhere but that a better understanding of community can be obtained by introducing network analysis, specifically ego-centered personal networks, which the book so well describes. Com-

mendably, the collection, perhaps unintentionally, does not neglect the cultural element that determine networks, a feature often unexamined by the formalizing trend usually preferred by many sociologists.

*Virtualism: A New Political Economy.* Edited by James G. Carrier and Daniel Miller. New York: Berg, 1998. Pp. v+222. $55.00 (cloth); $19.50 (paper).

Bruce G. Carruthers
*Northwestern University*

In his *General Theory of Employment, Interest, and Money* (Harcourt, Brace, Jovanovich, 1964), John Maynard Keynes worried that "practical men" had become the "slaves of some defunct economist" (p. 383). The argument of *Virtualism* is that we should be more concerned about the ideas of the living ones. In today's economy, abstraction is both an academic and a practical process. Economists construct mathematical models of the economy (a "virtual" economy), assuming or abstracting away the rich complexity and social embeddedness of economic life. But the economy is as reflexive as other social institutions, and so ideas about the economy react on the economy and change it. Economic models become the basis for public policy and private investment. The contributors see little virtue in virtualism and worry about the power of economic ideas. It is insufficient simply to note that these ideas are wrong, for they possess enough institutional power to "become" right.

James Carrier contributes two chapters to the collection, laying out the volume's agenda in his clear introduction. His other chapter addresses abstraction in economic practice and surveys economic changes in the West between the 18th and 20th centuries. He argues that, over this period, activities of both production and distribution became increasingly removed from the broader social contexts in which they had previously been embedded. His quick history of production covers familiar ground (cottage industry, through the putting-out system, and on to the factory), but readers will find his survey of retail trade interesting and valuable.

The purpose of Ben Fine's essay, "The Triumph of Economics," is to "examine both closely and critically some of the work of . . . Gary Becker" (p. 49). Becker is chosen as the exemplar of imperialist economics, which uses mathematical models, abstract to the point of dessication, to explain human behavior both in and outside the economy. Fine scores some telling points, but I would have been happier had his bibliography listed 10 works by Becker and a single work of Fine's, rather than the reverse. One worries that Becker functions as a virtual straw man.

Julie Nelson, who probably knows more neoclassical economics than any of the other contributors, offers a feminist critique of economics. She recreates a Sherry Ortner "female-is-to-male-as-nature-is-to-culture" style dichotomy and applies it to economic theorizing and then embraces (per-

haps without realizing it) Karl Polanyi by advocating a substantivist rather than formalist economics.

Helgason and Pálsson's chapter on Icelandic fishing quotas offers an insightful analysis of property rights, commodification, legitimacy, and the cultural limits of market exchange. Although fishing in Iceland may seem a tad narrow as a topic, readers can easily take their interesting ideas in several fruitful directions. Leslie Sklair sketches the globalizing equivalent of a power-elite argument: the transnational capitalist class dominates the global capitalist system using transnational corporations, and so on. Much of the analysis is predictable, even if probably true.

Philip McMichael examines the effect of structural adjustment on developmentalist states. Imposed by the International Monetary Fund (IMF) and the World Bank, structural adjustment programs force governments to curtail activist economic policies, cut spending, privatize public assets, and in general impose a "market logic" on public policy. Foreign debt served as the lever used to overturn developmentalist policies. McMichael's argument has the advantage of showing that "virtualist economics" is neither a disembodied mathematical canon, nor something hidden deep within the brain of Becker, but rather it is a set of formalized decision protocols and cognitive frames located within specific institutions (like the IMF).

Nigel Thrift's chapter points to other important institutional culprits: management consulting firms and business schools. Thrift is more interested in practical business knowledge and technique than haute economic theory and makes a good case for its significance. Management gurus like Tom Peters and mundane skills like bookkeeping matter more for global capitalism, Thrift claims, than do the latest economic theorems.

Daniel Miller's conclusion transforms notions of consumer sovereignty in an interesting way by arguing that consumption allows people to constitute diversity and idiosyncracy in the face of homogeneous and general global production. He also underscores the fact that economic doctrine is allowed the luxury of remaking the world in its own image and frequently does so in the name of the consumer. Economics is not the only academic discipline guilty of virtualism, however. Miller accuses postmodernists of doing essentially the same thing, albeit with less institutional clout than economics.

Overall, this volume contains an interesting mix of arguments and evidence. It raises some interesting questions but fails, I think, to explain adequately why economics is so privileged a discipline. How has it acquired the intellectual and institutional power to impose its abstractions on the real world? Or in other words, why do defunct sociologists not have more "practical men" as slaves?

*Pattern in Corporate Evolution.* By Neil M. Kay. New York: Oxford University Press, 1997. Pp. xvii+319. $65.00.

Christos N. Pitelis
*University of Cambridge*

In this important new book, Neil Kay deals with the nature, existence, and evolution of firms. He develops and employs the notions of *links* (which refer to the patterns of shared resources between elements of the economic system) and *patterns* (which refer to "consistency or balance" in the way the links are joined together). He builds on economic notions such as bounded rationality, scarce resources, opportunity cost, efficiency, and path dependence.

The book has 13 chapters, which may be usefully divided into four parts (not done by the author). The first part (six chapters) is introductory; it critiques existing approaches (chap. 2), introduces the notion of links and patterns (chap. 4), and reveals the framework (chap. 5). Kay does so in the context of firm strategies like vertical integration (chap. 3) or diversification (chap. 4).

The second part deals with firm *integration,* notably, diversification in chapter 7 and the multinational enterprise (MNE) in chapter 8. The third part deals with *co-operation;* joint ventures (chap. 9) and "alliances and networks" (chap. 10). In the last part, there is a return to issues of theory and a concluding chapter. The theory chapters deal with the link between hierarchy and complexity (chap. 11) and strategy and structure (chap. 12). Summary, conclusions, further research possibilities, and a reminder of the importance of "evolution" are discussed in the concluding chapter (chap. 13).

There are numerous important new insights here. Kay critiques both the mainstream *and* its critics. While the former deals with "optimal product-market price" (p. 9), critics have proposed only *partial* alternatives (satisfying instead of optimal, resource instead of product, hierarchy instead of market, and technology instead of price). Kay views "the firm as a hierarchically organized collection of resources making imperfect decisions in which technological change is typically the critical strategic variable" (p. 29).

On vertical integration, it is suggested that asset specificity (the key variable in Oliver Williamson's transaction costs agenda), fails to provide a satisfactory theory. *Asset replaceability* (and *appropriability*) "may provide a better guide as to why firms prefer to vertically integrate" (p. 56), especially in contexts other than production (e.g., research and development and advertising).

In the context of diversification, the major determinant of the boundaries of the firm is thought to be *resource commonality or linkage* across businesses. Links provide a source of synergy or internal economies and of potential shared vulnerability; thus, "the firms' choice of strategy may

be influenced by the balance and pattern of links associated with alternative strategies in different contexts" (p. 85). While richly linked strategies will normally be preferred, the threat of environmental attacks on particular links may encourage delinking.

To explain diversification, Kay proposes the notion of *coherence* to refer to strategies characterized by "linkages of *constant* strength and length" (p. 136; emphasis in original). Four strategies could display coherence: "the specialized, related-constrained, related-linked and conglomerate" (p. 150). Soft selection, path dependence, lock-in, and "turbulence" may help to explain the apparent stickiness of short-linked strategies.

The multinational enterprise (MNE) involves the internalization of specific and nonspecific assets. For Kay, "the MNE tends to be one of the weakest-linked options when analyzed in resource terms" (p. 156), "and should not be regarded as necessarily inevitable or inexorable" (p. 157). It is a sign of weakness, not strength.

Joint ventures (JVs) are seen as an expensive device for the coordination of resource allocation. They perform badly relative to single firm activity on power as well as efficiency criteria. The idea that collaborative activity may be triggered by (past) merger activity can contribute to an explanation of the recent proliferation of JVs. Alliances and networks could be attributable to the relative significance of the cooperative opportunities to the scale of the parties involved.

Other important claims of the book involve the idea that complexity of decision making generally diminishes as we move up the hierarchy. The use of one simple rule (that synergy dominates complementarity), Kay claims, can explain the logic of organizational structures.

In all, I am impressed but also must focus on limitations, which include a few interrelated issues: the employment contract, (endogenous) growth, dynamics, agency, oligopolistic interaction, and . . . Penrose. The first issue is hardly addressed, endogenous growth is hinted at but not utilized, the approach is in effect comparative static, it is not clear who are the agents of change (e.g., entrepreneurs, management?), oligopolistic interaction issues are extracted from, and Edith T. Penrose, arguably the mother-figure of the adopted resource-based perspective, is acknowledged but not fully utilized. In Penrose, internal (human) resources and resource indivisibility interact dynamically with the perceived external environment to determine productive opportunities. This leads to changes in both the internal and external environment and creates perceived opportunities for internal growth, mergers, vertical integration, and diversification. Resources, hierarchy, imperfect decisions, technology, information, and more are all here.

Penrose also suggests that multinationals too can be explained in these lines. Oligopolistic interaction and power considerations figure prominently in her work. It is unclear how Kay's analysis would be modified were the above features to be considered.

Despite limitations, this is a very important new contribution—a must-

read for researchers in this area. For students and managers it may be a hard read, as substantial knowledge of the subject matter seems to be presupposed.

*The Education-Jobs Gap: Underemployment or Economic Democracy.* By D. W. Livingstone. Boulder, Colo.: Westview Press, 1998. Pp. xv+341. $49.00.

*Created Unequal: The Crisis in American Pay.* By James K. Galbraith. New York: Free Press, 1998. Pp. xviii+350. $26.00.

William DiFazio
*St. John's University*

The United States in the 1990s has had seven years of economic boom with low unemployment, low inflation, and low government deficits. Amid all of this good news, inequality has increased and wages have barely risen. Commonsense knowledge seems to be right in this instance, that is, the rich get richer, the poor get poorer, and the middle class is shrinking. Though President Clinton boasts that the number of people on welfare has decreased significantly under his regime to 8 million, a 44% decline from 1994, he forgets that there are still 36.5 million poor people in the United States, which is only a 2% decline in the same amount of time. How is it possible that we have increasing inequality during economic prosperity?

This contradiction is not easily explained by the dominant neoclassical economic discourse of our time. Nor is it resolved by neoconservative social policy. More helpful are the two books under review: James K. Galbraith's *Created Unequal*, a Keynesian analysis of increasing wage inequality, and D. W. Livingstone's *Education-Jobs Gap: Underemployment or Economic Democracy*, a Marxist analysis of underemployment in both the United States and Canada.

James K. Galbraith provides a multicausal analysis that blames the current free market monetary policy for the increasing wage inequality. He calls for a rebellion in economic analysis and policy and for a reapplication of Keynesian macroeconomics to solve the problem. In *Created Unequal*, Galbraith successfully debunks the conservative contention that wage inequality is necessary because the new skill-biased technological innovation requires educated workers who are in short supply. For Galbraith, this is a fantasy. He also critiques their two other assertions: first, that global competition requires an increase in inequality and that the maintenance of inequality is necessary to fight inflation.

For Galbraith, the key to inequality is wage inequality that is politically determined, not the automatic result of the free market. He points to transfer payments that are mediated by the state: payments to the poor in the form of welfare are minor relative to payments to the elderly in

the form of social security or to the rich in the form of interest on public and private debt.

He argues that the evidence for the skill bias of new technology is sketchy at best. If anything, the evidence shows that there was skill bias in the 1960s and 1970s that did not continue into the 1980s. In fact, after 1984, we have essentially had continued wage stabilization. For Galbraith, what is most important about the new technologies is the monopoly power that they generate and the limitations they produce for the imputed perfect competition of the free market. It is in terms of technological monopoly that his agenda becomes clear. His main point is that wage inequality is the result of the economic policy of the neoclassical economists, especially Milton Friedman. Thus, the importance of monopoly power through the economy is that it sets prices, maintains wage inequality, and eliminates the rules of Friedman's competitive model.

Galbraith minimizes the social indicators of race, gender, and class and tells us that these are not important in understanding wage inequality. What is important is Keynesian macroeconomics. To make this point, he introduces a sectoral analysis of the economy. Here knowledge is dominant (the K-sector) and the producers of consumption goods (the C-sector) are in decline. The third sector is large and low paid (the S-sector). The K-sector controls the new technologies and wields monopoly power. Both wages and profits decline in the other two sectors. As a result of monopoly, power inequality increases.

The middle chapters of *Created Unequal* are the most powerful. Galbraith makes a very convincing argument that neoclassical monetary policy has resulted in increased wage inequality. "Inflation paranoia" is based on the belief that the natural rate of inflation must be higher than 5.5% to control inflation. For Milton Friedman, the nonaccelerating inflation rate of unemployment (NAIRU) requires an unequal society for a noninflationary economy. But Galbraith argues that historically the evidence does not back up the conservative argument, and unemployment does not predict the rate of inflation. The limits of NAIRU are exposed in the current economy that has low inflation and low unemployment. Galbraith thus argues for an alternative monetary policy. He calls it a comprehensive anti-inflation program without unemployment, with "low interest rates, high employment, a higher minimum wage supported by a stronger union movement, a maximum-minimum pay ratio and a national prospective inflation adjustment" (p. 245).

D. W. Livingstone in *The Education-Jobs Gap: Unemployment or Economic Democracy* would agree with Galbraith that it is a fantasy that the new technologies are skill biased. For him, there is not a shortage of highly skilled people. If anything, we are overeducated. The education-jobs gap, or underemployment, is the key social problem of our time: there are insufficient opportunities for our highly skilled workforce. For Livingstone, this is not just a problem for the United States but for Canada as well. In addition, it is a global problem, with 1 billion people unemployed or underemployed in the world (p. 2).

Highly skilled, university-educated workers have increasing problems finding decent jobs. The answer to the problem of underemployment has typically been to get more education, but Livingstone, armed with impressive evidence, shows that more education does not lead to better jobs. Education is not the answer; only a better economy can lead to better jobs.

The education-jobs gap has six dimensions: the talent-use gap, structural unemployment, involuntary reduced employment, the credential gap, the performance gap, and subjective underemployment. To this he adds a Marxist class analysis that demonstrates that underemployment is not just an economic accident, a result of the new technologies, or the result of human capital theory. It is the result of class-based struggles over education in capitalist societies. Thus, underemployment is the result of class-specific class conflict.

Not to deny the importance of Livingstone's analysis of underemployment, the most fascinating parts of this book are his analysis of the limitations of the unemployment rate, of human capital theory in terms of the six dimensions of unemployment, and his critique of the postindustrial knowledge society's requirement of higher skill. He shows that, in spite of the creation of higher and lower skilled work, there are still not enough jobs that pay well. Often the problem is that more education is required for the same jobs, which results in credential inflation. He argues that tacit, informal, on-the-job education is always important but rarely recognized and that this component should be taken more seriously. What makes this book so compelling is his excellent use of the quotes of his informants from extensive qualitative interviews.

Both books suffer from an inability to make their powerful arguments into a political program. I believe that part of the problem is the way that both authors view history as a continuous process that enframes the limits of their analysis but does not allow for the possibility of social change. An example of this is their narrow definition of technology. They see technology as a continuum, thus computers and robotics are just another lever. But if the new technologies are universal, as Alan Turing said they were, they can replace not only physical function but knowledge function in the workplace as well. Then, instead of historical continuity, we would have a break in history, with part of the reason for underemployment and wage inequality residing in the specific global use of these labor-destroying technologies to create new power bases. Of course, the new technologies might be the lever to create new globally integrated modes of work and communication which could result in a more equal and democratic society. To conclude, both Galbraith and Livingstone have written important books that, in complementary ways, add not only to the discourse of ending inequality but provide us with new and important ways to understand the problem.

*Parting at the Crossroads: The Emergence of Health Insurance in the United States and Canada.* By Antonia Maioni. Princeton, N.J.: Princeton University Press, 1998. Pp. xiv+205. $37.50.

Brian Gran
*Yale University*

For health insurance, the United States has taken the road less traveled. The United States is the only rich country without universal health insurance. People in the United States spend the most, rely heavily on the private sector, and obtain care from the world's most complicated delivery system. While some supporters have expressed satisfaction, if not pride, in these remarkable qualities, others contend that the United States faces *unique* limitations in reforming health care.

In her exceptional book, *Parting at the Crossroads*, Antonia Maioni compares the formation of the U.S. and Canadian health-care systems for the years 1930–60. Two of the most significant contributions Maioni makes are to demonstrate which political institutions mattered for how political actors succeeded or failed at health reform and, in particular, how federalism, often considered an impediment to policy formation, was key to Canadian reforms.

The United States and Canada are often considered the most similar of Western democracies. They share a common border, are wealthy, and have federal governments. Their trade unions are only moderately powerful, and their populations are diverse and young. Nevertheless, their health-insurance systems are nearly opposite. The United States relies on a mix of government plans, targeted to the elderly and indigent, and employment-based plans, which the government indirectly supports. Canada offers public health insurance to all qualified residents, with the private sector providing supplementary services in some provinces.

Maioni demonstrates that neither the United States nor Canada had any major initiative in health policy prior to the 1930s. With the Depression came opportunities for introducing public health insurance. The U.S. Social Security Act of 1935 laid the foundation for the U.S. welfare state, but politics hampered health-insurance initiatives. In contrast, the Canadian national government hesitated on social policy reform because of constitutional questions. The U.S. social-insurance approach established a policy legacy, but policy absence meant a clean slate for Canadian health-policy innovation.

Labor organizations became strong advocates for health-insurance reform in both countries. Their impact partially depended on political institutions and how other actors, particularly organized medicine, wielded them. Canada's governmental and electoral systems allowed labor to cooperate with a social democratic party in the Saskatchewan province, which established a universal program. The Saskatchewan program demonstrated universal insurance feasibility, spurring the dominant Liberals to introduce a national universal program. In contrast, the U.S. electoral

system effectively precluded third-party formation, forcing organized labor to dilute its health-insurance goals because it was one of many interests represented by the Democratic Party.

Organized medicine in both countries battled reform efforts. The American Medical Association, taking advantage of public wariness about government insurance, lobbied Congressional members, gaining bipartisan support for its voluntary alternative. Employing the Social Security legacy, organized labor responded with insurance for the aged, which became Medicare. Although Canadian organized medicine opposed national health insurance, its efforts were hampered by political party discipline. Once the governing party set the agenda, the Canadian Medical Association could not influence individual political actors like U.S. organized medicine did.

Maioni suggests that economic vitality is important for the future of both countries' systems, but the prognosis is uncertain. Despite recent concerns about the Canadian government's budgetary health, Maioni contends that widespread support protects universal insurance. Conversely, Maioni seems pessimistic about options for U.S. universal health insurance. Despite economic buoyancy, dissension will likely prevent reforms. Although a devastating economic downturn would make health finance difficult in either country, the U.S. system seems especially vulnerable. Employment-based insurance and Medicare both rely on labor market attachment. High, chronic unemployment could result in coverage loss and financial difficulties for employer insurance and Medicare, swelling the uninsured pool. Such a crisis could provide an opening for universal health insurance. In any case, whether the United States relies on the public or private sector, escalating health expenditures figure into budgets of governments, corporations, and families. The U.S. health care system's future may depend on Americans' willingness to devote more of their national income to health care.

Through her book, *Parting at the Crossroads*, Antonia Maioni makes an outstanding contribution to the study of health insurance and the welfare state. Maioni demonstrates that, despite similar starting points, political institutions, and actors, Canada successfully introduced, while the United States continues to resist, universal health insurance. Students of political sociology, health policy, and the welfare state, and policy analysts and makers will benefit from reading Maioni's impressive book.

*Political Parties, Growth and Equality: Conservative and Social Demo-
cratic Economic Strategies in the World Economy.* By Charles Boix. New
York: Cambridge University Press, 1998. Pp. xiv+280. $59.95 (cloth);
$19.95 (paper).

Thomas Janoski
*University of Kentucky*

Boix states that "most scholars—as well as many voters—question
whether electoral politics and political parties have any significant impact
on public policy, especially policies that are concerned with economic
growth and full employment" (p. 2). To demonstrate that parties have a
key impact on economic policy, he divides his book into two parts. In
part 1, he develops the theoretical approaches of conservative and social
democratic parties using the supply-side economics of human capital, in-
vestment, and taxes and then tests some of these propositions against
data in 20 countries. In part 2, Boix develops case studies of the social
democratic project of the government in Spain and the conservative strat-
egy of the Thatcher regime in the United Kingdom. The two approaches
are largely parallel in demonstrating his claim that parties matter more
than "most scholars" think they do.

In the theoretical core of the second chapter, Boix graphs the distribu-
tion of skills, from unskilled to very skilled, and the percentage distribu-
tion of the population. In spite of an occasionally confusing introduction,
Boix does effectively delineate the growth strategies of conservatives and
social democratic regimes. The conservative strategy is to lower the social
wage (unemployment compensation) in line with the declining skill levels
of the unemployed and to keep taxes low on the highly skilled in order
to promote private investment. The social democratic strategy is to raise
the skills and increase the wages of the unemployed through public in-
vestment in education and training but at the same time to raise taxes
and reduce the incomes of the most highly skilled. Both strategies can
be successful in producing economic growth, but each one encounters its
own stresses and pitfalls. Boix then connects the lower half of the skill
and income spectrum with the social democratic voters and the higher
half with the conservative voters. The battle is over the volatility at the
center, especially when economic shocks disrupt the economy. Centrifu-
gal pressures also exist at the extremes with low-income voters being at-
tracted to protectionist or radical parties on the left and higher-income
voters taking up with libertarian parties on the right.

In his empirical tests, Boix shows that socialist governments increase
the money spent over GDP on education, active labor market policies,
and gross fixed capital investment. Socialist governments tend to increase
the tax rates on high-income people but avoid privatization of govern-
ment industries. Using conservative party power with tax reductions and
privatizations would presumably produce opposite results. This all helps

to prove his claim that parties matter for policy selection, though he does not want to go so far as to claim influence on economic outcomes.

The case studies generally support the statistical evidence. The Gonzàlez government in Spain was able to create a social democratic coalition that pursued human capital policies, resisted demand stimulation policies, and kept taxes high on the upper classes. The Thatcher government created a conservative coalition that cut some unemployment benefits but kept the health service and pensions somewhat sacrosanct, decreased taxes, privatized businesses and housing, and generally relied on private-sector investment. Both governments were moderately successful, according to Boix, but both coalitions faced tensions, concessions, and difficult trade-offs.

The strength of this book is its examination of investment, tax, and privatization policies in combination with education, social, and active labor market policies. The most innovative part of the book is how changing policy mixes can create new electoral blocks based on reshaped class configurations. In Spain, low-income rural areas are converted to the social democratic fold. But most interesting is how the Thatcher government sold off the government's tenant housing to convert labor-minded renters into conservative-supporting home owners. The simulations of policy choices under different social conditions also present innovative ways of viewing counterfactual situations. Not being apparently partisan, Boix fairly evaluates conservative and social democratic policies in their complete array of differing emphases. On the whole, Boix is successful in showing that parties matter in choosing different policy strategies, and even though he does not specifically claim it, those policies have an important impact on society and the economy.

To get at many of these valuable insights, however, the reader has to put up with some annoying barriers. Boix is given to overstatement (e.g., all, always, dominant, etc.) followed by few references. For instance, few would agree with the quote that begins this review, stating that "most scholars" would question that political parties have any impact. Just for starters, the many sociologists and political scientists who have argued for the impact of left party power and used power resources theory would certainly be surprised at his statement and his neglect of their work. Further, we never find out who "most scholars" are. Unfortunately, Boix takes a wide-ranging view of the policy field with a very narrow focus on the literature. Presentation is also a problem. The five full-page figures tracing policy expenditures in 20 countries over time force the reader to connect the dots and then try to make sense of the results. Tables with countries grouped according to growth rates in expenditures would have taken up less space and have been much more comprehensible. And unless one is demonstrating important outliers, graphing correlation coefficients is not necessary. Finally, the later sections on the case studies are

somewhat repetitive despite avoiding nearly all references to actors (save Gonzàles and Thatcher and precious few others), and the penultimate chapter should be folded into the conclusion. Nonetheless, for those who defeat these barriers, the book provides a complex and coherent account of political party strategies and the impact they can have on the electorate and on society as a whole.